SYMMETRY OF CURVES AND THE GEOMETRY OF SURFACES: TWO EXPLORATIONS WITH THE AID OF COMPUTER GRAPHICS.

Thesis submitted in accordance with the requirement of the University of Liverpool for the degree of Doctor of Philosophy

by

Richard James Morris,

May 1990.

PAGE NUMBERING AS ORIGINAL

VOLUME CONTAINS CLEAR OVERLAYS

OVERLAYS HAVE BEEN SCANNED SEPERATELY AND THEN AGAIN OVER THE RELEVANT PAGE

Dedication

To the planet, in the hope that this work will be of use in the future.

Acknowledgements

To my supervisor Dr. P. J. Giblin, for introducing me to the wonderful world of singularities. To my room mate Farid Tari for many interesting discussions (not all about mathematics). To the University Green Group and associates, for taking my mind of the mathematics (occasionally a little to often). To the Science and Engineering Research Council and the Department of Social Security for financial support.

Abstract

This thesis consists of two main parts: the local symmetries of plane curves and the geometry of surfaces. In each part we develope algorithms for finding various sets and show some computer generated examples.

In the first part we study the Rotational Symmetry Set, the Symmetry Set and the Mid-Point Locus. These sets contain information about the local symmetries of a plane curves γ. For each point on one of these sets there will be either a rotational or a reflectional symmetry between two points on γ. Examples of most of the transitions involving these sets have been calculated.

For any smooth surface S we can calculate the focal surface F which consists of all the centres of spheres which have A_2 or higher contact with S. In the second part we study the cuspidal edges (ribs) and parabolic lines on the focal surface F as well as the corresponding ridges and sub-parabolic lines on the surface S. In particular we look at the relationship with the Gaussian curvature of the focal surface. Numerous examples are shown of the transitions of the patterns of ridges and sub-parabolic lines on the surface. We also look at three dimensional pictures of the main types of umbilic focal surface.

Table of Contents

Chapter 1 Introduction 1

1.1 Local symmetry of curves 1
1.2 Surfaces 4

PART 1 SYMMETRY OF CURVES

Chapter 2 Defining the Symmetry Set, Rotational Symmetry Set & Mid-point Locus 8

2.1 Contact of curves 8
2.2 The centre maps and the Rotational symmetry Set 10
2.3 Defining the Symmetry Set and Mid-point Locus 13
2.4 The local structure of the Rotational Symmetry Set 15
2.5 The local structure of the Symmetry Set and Mid-point Locus 20

Chapter 3 Calculating the Rotational Symmetry Set 25

3.1 Summary of algorithm 25
3.2 The movement of the points 26
3.3 Dealing with two extrema of equal y-value 29
3.4 A different way of regarding the problem 30
3.5 Have all the pairs (L,R) been found? 33
3.6 Have any pairs been found twice? 40
3.7 Finding the RSS part 2 41

Chapter 4 Calculating the Symmetry Set 42

4.1 A function h which defines the Symmetry Set 42
4.2 Following the $h^{-1}(0)$ contour 46
4.3 Examples of the operation of the routine 54
4.4 Dealing with orientation 55
4.5 Finding starting points 57

Chapter 5 Examples of transitions on the Rotational Symmetry Set, Symmetry Set and Mid-point Locus 64

5.0 Key to Examples 66
5.1 Generating Examples 67
5.2 The κ–κ' diagram 69
5.3 The A_4 Transition on the RSS 71
5.4 The swallowtail transition 73
5.5 The I-Lips transition 73
5.6 Two equal maxima of curvature 76
5.7 The I-Beak transition 79
5.8 The trident transition 81
5.9 Triple crossing on the Rotational Symmetry Set 83
5.10 Events involving type II inflections on RSS and SS 85
5.11 The A_4 transition: RSS, SS and MPL 87
5.12 The biosculating case with $\kappa'(t_1)\kappa'(t_2)>0$ (moth transition on SS) 89
5.13 The biosculating case with $\kappa'(t_1)\kappa'(t_2)<0$ (nib transition on SS) 89
5.14 A cusp on the Mid-point Locus 89
5.15 Transitions on the second part of the RSS 94
5.16 The κ–κ' diagram for the RSS part 2 97
5.17 The Half Lips transition on the RSS part 2 98
5.18 The Half cusp transition 99
5.19 Duals of transitions 101
5.20 Piecewise circular curves 103

PART 2: SURFACES

Chapter 6 Introduction to the geometry of surfaces 111

6.1 Surfaces and the derivatives of maps 111
6.2 The first and second fundamental forms, the Gauss map and principal directions 113
6.3 Contact with spheres and the distance squared function 115
6.4 Focal surface and their singularities 117

Chapter 7 Calculations on focal surfaces 120

7.1 The relationship between the derivatives of the distance squared 120

map and its singularities
7.2 Calculating the Gaussian curvature of the focal surface 123
7.3 A simplified form of the Gaussian curvature of the focal surface 126
7.4 Parabolic lines and cuspidal edges on focal surfaces 129

Chapter 8 Examples of Umbilics 136

8.1 Cubic forms 136
8.2 Umbilic classification 140
8.3 Generating examples of umbilics 151
8.4 Orthogonal transition 152
8.5 Parabolic umbilic transition 154
8.6 Birth of umbilics 154
8.7 Monstar - Lemon transition 157
8.8 Symmetric umbilics 157
8.9 Examples of focal surface 157
8.10 Parabolic lines on an elliptic umbilic focal surface 162
8.11 Parabolic umbilic transition 166
8.12 Hyperbolic umbilic focal surfaces 178

Chapter 9 The program 185

9.1 Representing the surface 185
9.2 Finding features on the surface 186
9.3 The zero following routine 187
9.4 Finding zeros 188
9.5 Finding starting points 191
9.6 Processing umbilics 192
9.7 Finding where two features cross 194
9.8 Introduction to drawing surfaces 195
9.9 Drawing the image of a curve in parameter space 199
9.10 Rectangular grids 199
9.11 Polar grids 201
9.12 Creating *nets* and *tmeshes* 205
9.13 Other methods 208

References 211

Chapter 1: Introduction

This thesis consists of two main part. The first part concerns the local symmetry of plane curves and the second concerns the geometry of surfaces. In each part we calculate various examples with the aid of a computer. In the main text when referring to a figure generated by computer we will use a $*$, for example figure 1.6)* has been generated by a computer.

§1.1 Part 1 Local Symmetry of Curves

Given any two points $\gamma(t_1)$, $\gamma(t_2)$, on a parametrized curve γ we have two points $C_1(t_1,t_2)$, $C_2(t_1,t_2)$ defined by the *Centre Maps* (Def. 2.2.1). About these points there exist rotations which send $\gamma(t_1)$ to $\gamma(t_2)$ and the tangent line to γ at $\gamma(t_1)$ to

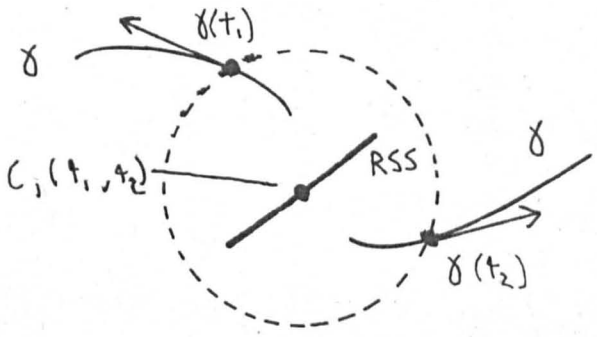


fig. 1.1) A point on the Rotational Symmetry Set.

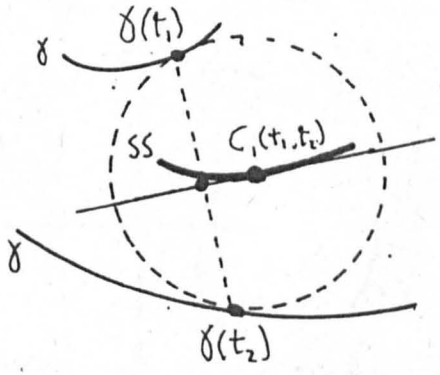


fig. 1. 2) Points on the Symmetry Set and Mid-Point Locus.

the tangent line to γ at $\gamma(t_2)$ (fig. 1.1). The set of pairs (t_1,t_2) can be restricted so that the corresponding set of centres contain information about the local symmetry of the curve. If we choose t_1, t_2 so that the curvatures at the two points are equal

in magnitude, then there will be a high degree of local rotational symmetry and the centres form the *Rotational Symmetry Set* **(RSS)**. If on the other hand we consider those pairs for which a circle centre $C_1(t_1, t_2)$ is tangent to γ at both $\gamma(t_1)$ and $\gamma(t_2)$, (a *Bitangent Circle*), then there will be an axis of local reflectional symmetry which passes through $C_1(t_1, t_2)$ and $\frac{1}{2}(\gamma(t_1) + \gamma(t_2))$ (fig. 1.2). We call the set of centres of such pairs the *Symmetry Set* **(SS)** and the midpoints of the chords joining $\gamma(t_1)$ and $\gamma(t_2)$ the *Mid Point Locus* **(MPL)**.

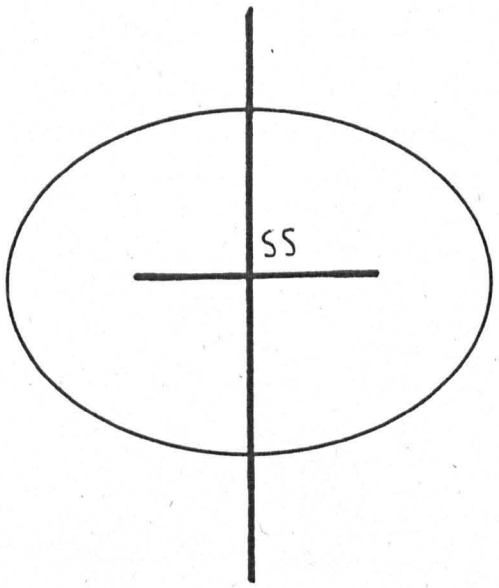


fig. 1.3a) The Symmetry Set of an ellipse.

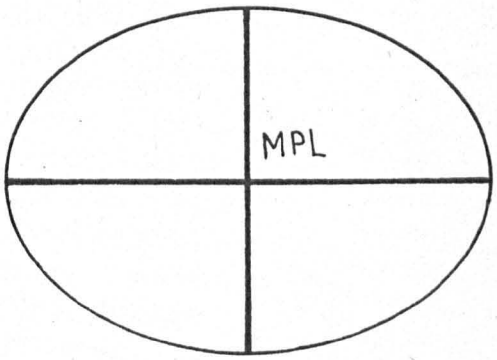


fig. 1.3b) The Mid-Point Locus of an ellipse.

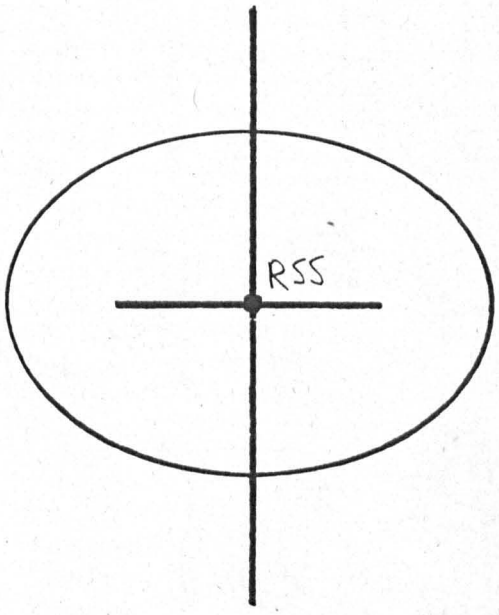


fig. 1.4) The Rotational Symmetry Set of an ellipse.

For highly symmetrical objects such as an ellipse the SS and MPL actually coincide with the axes of symmetry (fig. 1.3). In this example the Rotational Symmetry set is highly concentrated in the centre of the ellipse (fig. 1.4) rotation by 180° about this point will map the ellipse onto itself.

The special points in these sets (cusps, inflections, points at infinity and end points) usually correspond to higher symmetries. For example when the RSS and SS both have inflections at the same point there is both local rotational and reflectional symmetries.

The study of Symmetry Sets started with Blum's symm-axis [Blum], which was used as a way of characterizing shapes of biological objects (fig. 1.5). By finding the centres of all bitangent circles contained entirely inside an object a representation of the object can be found.

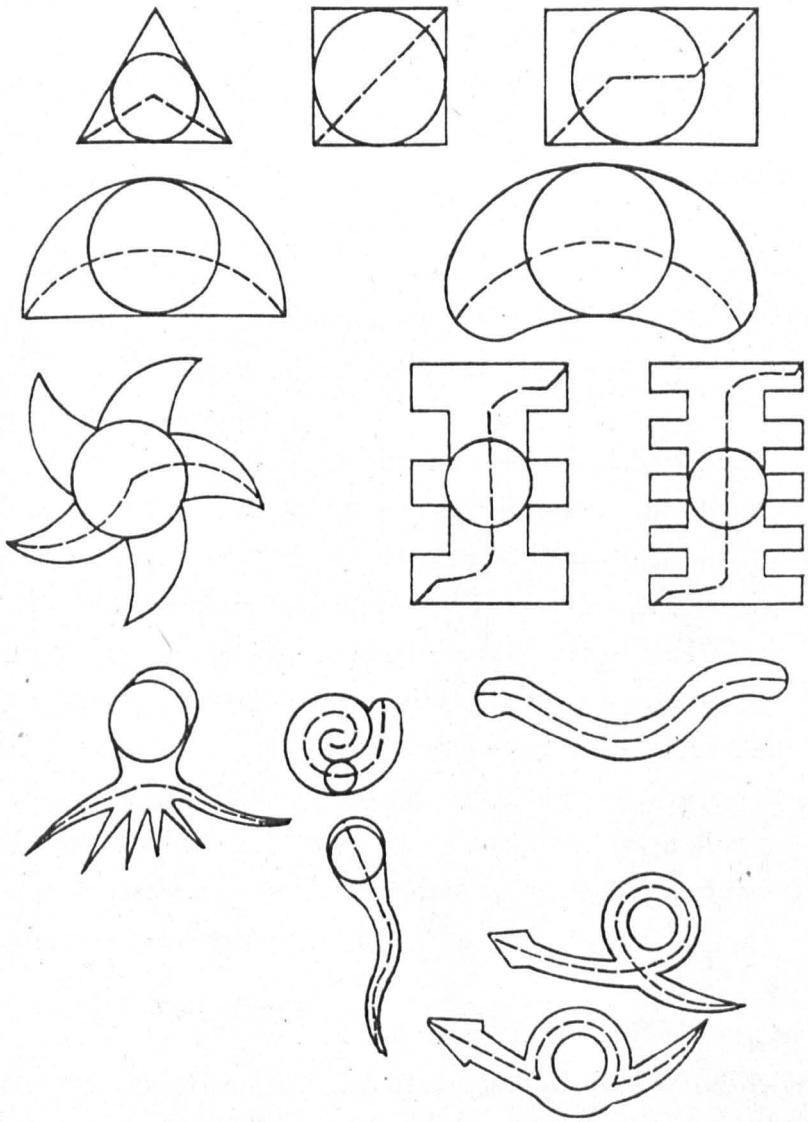


fig. 1.5) Some examples of the Symm-Axis (Reproduced from [Blum] fig. 1.8).

The symmetry set is particularly interesting in the field of robot/computer vision [Brady] as it provides a means of simplifying the information stored about an object. There are also implications in the area of human vision [Leyton] and robot path planning [Canny].

Several attempts have been made at calculating the SS or MPL [Scott–Turner–Zisserman] (SS) [Brady–Asada] (MPL). These have concentrated on finding the Symmetry Sets of real world objects where incomplete information about the object is known. While good for the intended objective the fine detail of the Symmetry Set is missed by these algorithms. The algorithm presented here requires much more information (up to the third derivative) about the object but produces very detailed results.

The mathematical background behind the Symmetry Set has been thoroughly studied [Bruce-Giblin] [Bruce-Giblin-Gibson] [Giblin-Brassett] [Giblin-Tari] and [Tari] and all the generic local forms are known. However a complete set of pictures of these local forms produced from actual curves has not been obtained. The pictures in this work rectify this omission and have helped answer some questions particularly about the dual structure.

The Rotational Symmetry Set is the first attempt at generalizing the Symmetry Set to cover local rotational symmetry. As well as possible geometrical interest in its own right the RSS also occurs as part of the critical set of the Centre Maps so is interesting from a mathematical point of view. The pictures here first highlighted the similarity between the dual of the Symmetry Set and the Rotational Symmetry Set [Giblin-Tari].

Chapter 2 introduces the mathematical background for these sets. The two algorithms for finding the sets are discussed in chapters 3 and 4. In chapter 5 examples are provided of most transitions which can occur as well as their duals and the Symmetry Sets of piecewise circular curves. In this chapter a method of quickly predicting the topological structure of the RSS is introduced.

Throughout the first half of the project I worked closely with Farid Tari. While Farid was mainly concerned with the mathematical analysis of the various cases I have concentrated on the computing. There have been many instances where we have worked on the same idea simultaneously. Sometimes the idea arose from an example generated by the computer and at other times the computer was used to confirm a previously established result.

§1.2 Part 2 Surfaces

The second part of this thesis is devoted to finding smooth surfaces and their associated focal surfaces. At each point of a surface there are two principal directions

and two corresponding principal curvatures κ_p, κ_q. The *Focal Surface* consists of two sheets, one for each of the principal curvatures (fig 1.6). For each point on the surface the focal points lie distances $1/\kappa_p$ and $1/\kappa_q$ along the normal to the surface.

As well as finding the surfaces we also calculate various features on them. For example we will find the set of local extrema of each of the principal curvatures

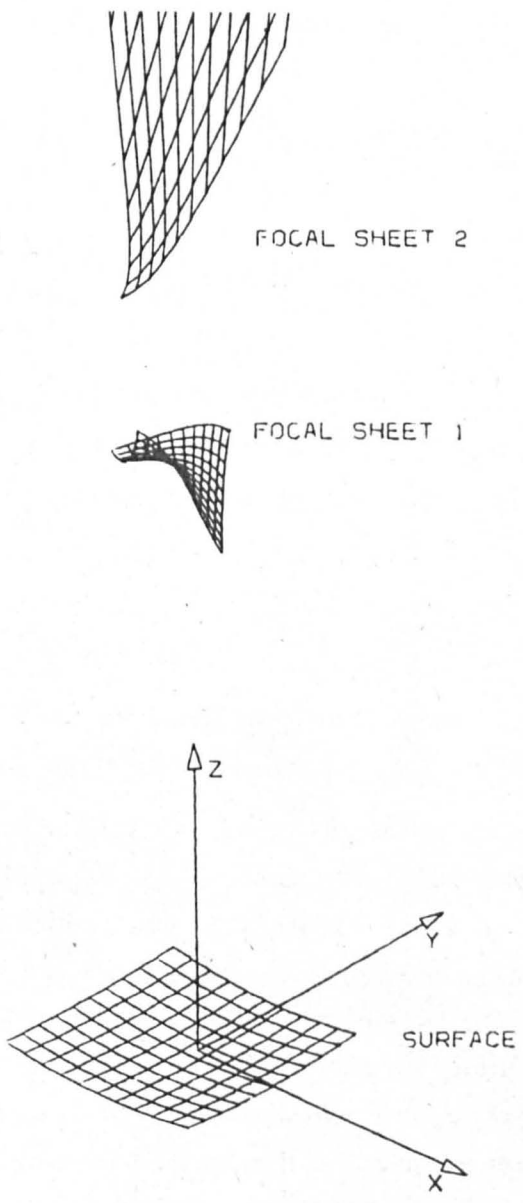


fig. 1.6)* A surface and its focal sheets.

(or more precisely A_3 points) these form the *Ridges* of the surface. On the focal surface the corresponding point will form cuspidal edges. An other feature that we will study are the parabolic lines on the focal surface (which give rise to sub-

parabolic lines on the original surface). Algorithms are presented here for finding both of these sets on a generic parametrized smooth surface.

When both principal curvatures are equal we have very special points called *Umbilics*. These umbilics will generically be isolated points lying on a ridge. In fact we would normally expect either 1 or 3 ridges and 1 or 3 sub-parabolic lines to pass through each umbilic [Bruce-Wilkinson] [Porteous]. The focal surface of such points is highly singular with both sheets of the focal surface coming together at a single point. There are several interesting areas of study relating to these features, as follows.

1. The pattern of ridges and sub-parabolic lines on the surface (or equivalently in the parameter space).
2. The connection between sub-parabolic lines and the pattern of lines of curvature. For example when changing from "Monstar" to "Star" patterns the number of sub-parabolic lines through the umbilic changes from 3 to 1.
3. The relationship between parabolic lines, cuspidal edges and the Gaussian curvature of the focal surface. We shall see later that the Gaussian curvature will generically change sign when we cross a cuspidal edge.
4. Obtaining pictures of focal surfaces and in particular the highly singular points such as umbilics and swallowtail points. Some progress has even been made in looking at families of focal surfaces.
5. Studying the ridges and sub-parabolic lines on the surface to see if they correspond to visually significant features.

Here we develope a program for calculating ridges and sub-parabolic lines and a program for drawing focal surfaces. Together with some mathematical analysis these programs have been used to study 1), 2), 3) and 4) above. The program could be used to study 5) as well, but this has not be carried out as the examples for this have different characteristics to those of the other four parts.

The mathematical background and classification of points on the surface is discussed in chapter 6. This is elaborated in chapter 7 where we describe functions for finding ridges and sub-parabolic lines as well as making a study of the Gaussian curvature of focal surface. In chapter 8 we discuss the classification of umbilics and show examples of the transitions which can occur on both the surface and focal surface. An algorithm for finding the various features on a parametrized surface patch is discussed in chapter 9. Also chapter 9 deals with the problems of drawing the highly singular focal surfaces.

PART 1

SYMMETRY OF CURVES

Chapter 2: Defining The Symmetry Set, Rotational Symmetry Set & Mid-Point Locus

In this chapter we define the Rotational Symmetry Set (RSS), the Symmetry Set (SS), and the Mid-Point Locus (MPL) (§2.2, §2.3), the concept of contact between (§2.1) two curves and the "centre maps" which will be useful tools later. The local structure of the sets is described in §2.4 and §2.5. The chapter ends with an example of the various sets (§2.6). A full mathematical analysis of these sets can be found in [Tari].

§2.1 Contact of curves

In this section we review some standard results about the contact of curves. For more details see [Bruce-Giblin-2].

The curve $f(t) = (x(t), y(t))$, which is parametrized by arc length, and the curve $G^{-1}(0)$, where G is a function from the plane to the real line, have $n+1$ **point contact** at $f(t_0)$ if

$$(G \circ f)^{(k)}(t_0) = 0 \quad for\ 0 \leq k \leq n,$$
$$(G \circ f)^{(n+1)}(t_0) \neq 0.$$

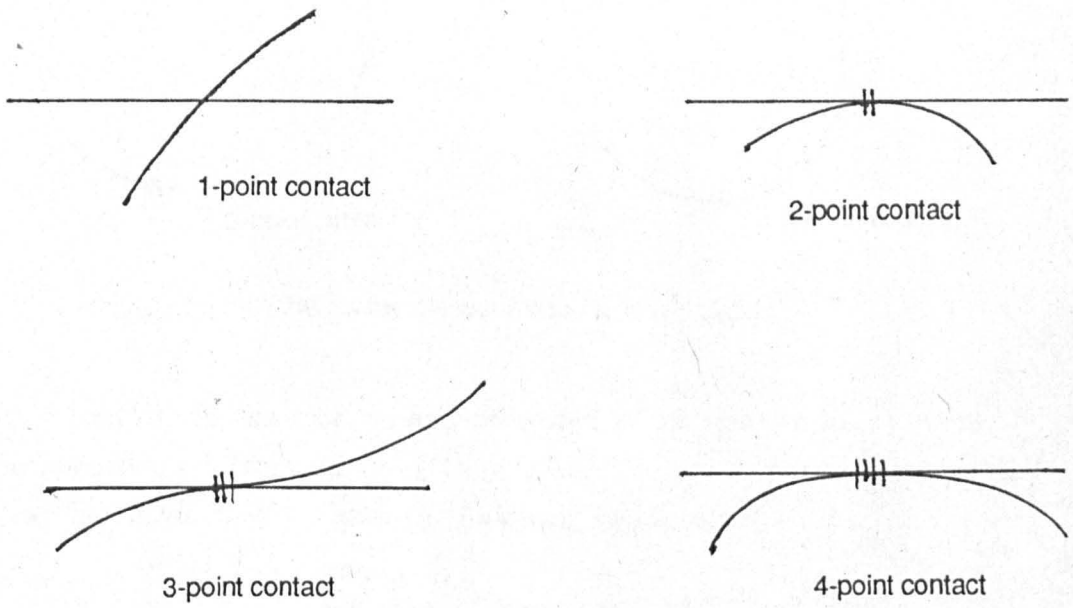


fig. 2.1) The contact between a curve and a straight line.

For example if $G(x, y) = x$ the $G^{-1}(0)$ is the y-axis and the order of contact

depends on how many derivatives of $x(t)$ vanish. There will be 1-point contact if $f(t)$ crosses the x-axis transversely, 2-point or higher contact if $f(t)$ is tangent to the x-axis ($x'(t) = 0$), 3-point contact if $f(t)$ has an simple inflection at the point of contact ($x''(t) = 0$) and 4-point or higher if $f(t)$ has a higher inflection ($x'''(t) = 0$) (fig. 2.1). Now consider a circle of radius r centre the origin which is given by $G(x, y) = x^2 + y^2 - r^2$. The curve $f(t)$ will have 1-point contact with $G^{-1}(0)$ if $f(t)$ crosses the circle transversely, 2-point or higher if $f(t)$ is tangent to the circle, 3-point or higher if the curvature, $\kappa(t)$, at the point of contact is $\frac{1}{r}$, 4-point contact if in addition $\kappa'(t) = 0$ (fig 2.2). Higher contacts depend on how many derivatives

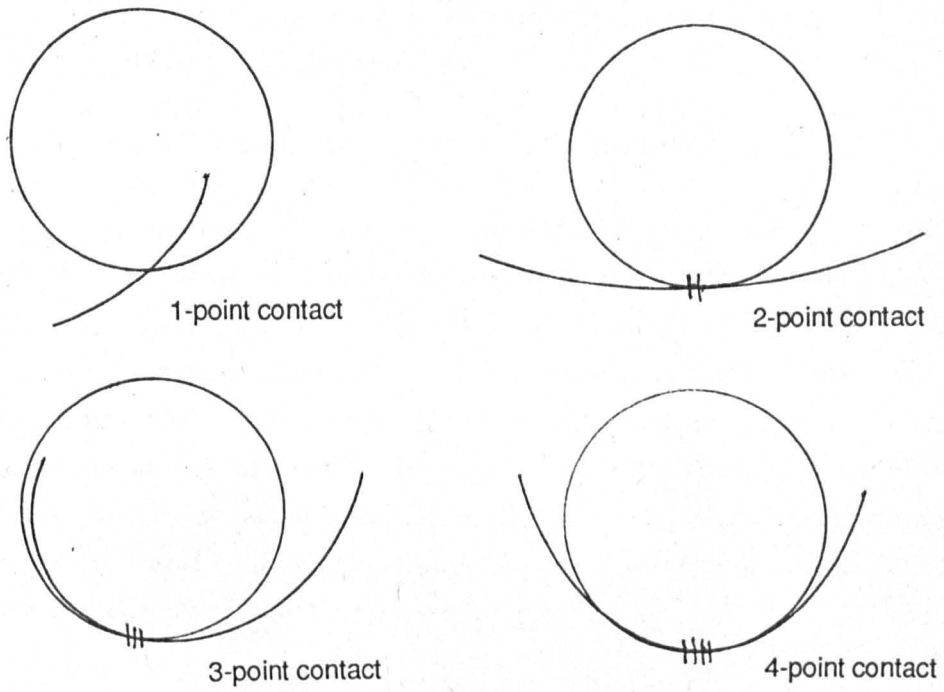


fig. 2.2) The contact between a curve and a circle.

of curvature vanish. In the case we are interested in we want to know about the contact between two arbitrary curves $f(t)$, $g(t)$ which are both parametrized by arc length it can be shown that we have the following conditions.

1-point contact if $f(0) = g(0)$, $f'(0) \neq g'(0)$.
2-point contact if $f(0) = g(0)$, $f'(0) = g'(0)$, $\kappa_f(0) \neq \kappa_g(0)$.
3-point contact if $f(0) = g(0)$, $f'(0) = g'(0)$, $\kappa_f(0) = \kappa_g(0)$, $\kappa'_f(0) \neq \kappa'_g(0)$.
4-point contact if $f(0) = g(0)$, $f'(0) = g'(0)$, $\kappa_f(0) = \kappa_g(0)$, $\kappa'_f(0) = \kappa'_g(0)$, $\kappa''_f(0) \neq \kappa''_g(0)$.

We will also use the A_k notation for contact of curves with circles.

A_1 contact $\iff$ 2-point contact $\Rightarrow$ $f(t)$ is tangent to the circle.
A_2 contact $\iff$ 3-point contact $\Rightarrow$ the circle is the **Osculating Circle** at point of contact.
A_3 contact $\iff$ 4-point contact $\Rightarrow$ $f(t)$ has a vertex at point of contact
A_4 contact $\iff$ 5-point contact $\Rightarrow$ $f(t)$ has a higher vertex at point of contact

We can extend this notation to cover circles which touch the curve in more than one place. We say the circle has $A_m A_n$ contact with the curve if it has A_m contact in one place and A_n contact in another place. In particular a bitangent circle has $A_1^2 = A_1 A_1$ contact with the curve.

§2.2 The Centre Maps and the Rotational Symmetry Set

We now return to the problem of defining the three sets. Let $\gamma_i : I \to \mathbb{R}^2 \{i = 1, 2\}$ be two parametrizations of unit speed smooth plane curves. Here I is either the unit interval or for closed curves the circle S^1. Also let $\kappa_i(t)$ be the curvature, $T_i(t)$ be the unit tangent vector and $N_i(t)$ be the unit normal at the point $\gamma_i(t)$. For the rotational symmetry set we wish to find the set of points $\{x\}$ in the plane about which there is a rotation $R_x : \mathbb{R}^2 \to \mathbb{R}^2$ such that there is 3-point contact between $R_x(\gamma_1)$ and γ_2 at some point $R_x(\gamma_1(t_1)) = \gamma_2(t_2)$. For 2-point contact we require that $R_x(\gamma_1)$ and γ_2 have the same tangent line at this point. A rotation about a point, x, through angle θ can be expressed as

$$R_x : z \to (z - x)e^{i\theta} + x$$

where $\mathbb{R}^2$ is identified with $\mathbb{C}$ and multiplication is as for complex numbers. The derivative of this map is

$$R_x' : T \to Te^{i\theta}$$

where T is some tangent vector which we think of as a complex number. So for 2-point contact we require

$$\gamma_2(t_2) = (\gamma_1(t_1) - x)e^{i\theta} + x$$

and

$$T_2(t_2) = \pm T_1(t_1)e^{i\theta}.$$

Now

$$e^{i\theta} = \pm T_2(t_2)/T_1(t_1)$$

so by substitution and rearrangement we have

$$(T_1(t_1) \mp T_2(t_2))x = \mp T_2(t_2)\gamma_1(t_1) + T_1(t_1)\gamma_2(t_2).$$

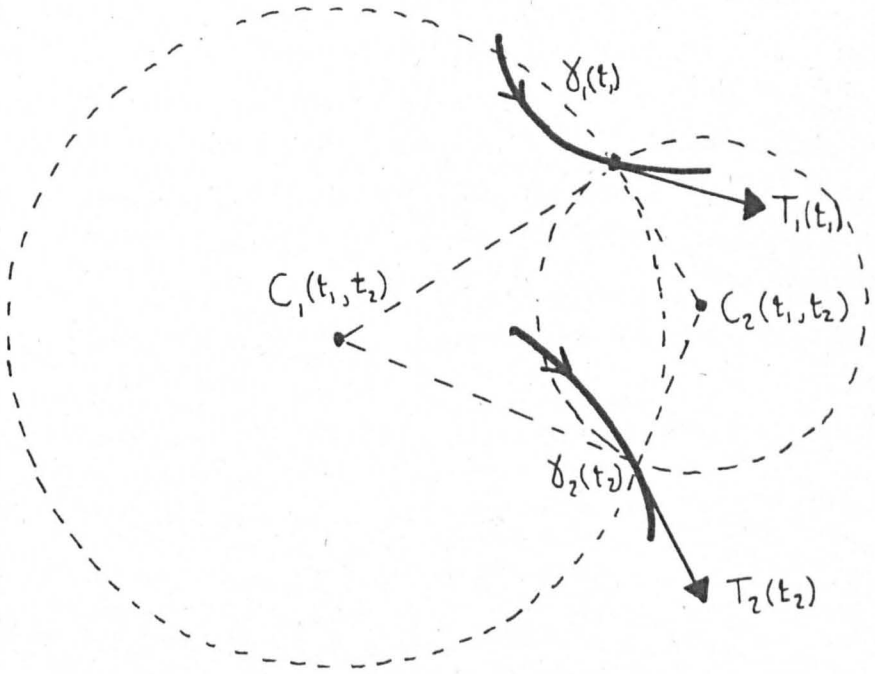


fig. 2.3) The two centre maps.

Definition 2.2.1

The **first centre map** $C_1 : I \times I \to \mathbb{R}^2$ is defined by

$$C_1(t_1, t_2) = \frac{\gamma_1(t_1)T_2(t_2) - \gamma_2(t_2)T_1(t_1)}{T_2(t_2) - T_1(t_1)}$$

where $\mathbb{R}^2$ is identified with $\mathbb{C}$ and multiplication and division happen as for complex numbers. This map is defined for all t_1, t_2 such that $T_1(t_1) \neq T_2(t_2)$. There is a rotation R_1 about this point such that $R_1(\gamma_1(t_1)) = \gamma_2(t_2)$ and $R_1'(T_1(t_1)) = T_2(t_2)$ (fig. 2.3). If $\gamma_1 = \gamma_2$ then $C_1(t_1, t_2)$ has a limit as $t_1 \to t_2$ which can be found by L'Hopital's Rule,

$$\begin{aligned}
\lim_{t_1 \to t_2} C_1(t_1, t_2) &= \lim_{t_1 \to t_2} \frac{\gamma_1(t_1)T_2(t_2) - \gamma_2(t_2)T_1(t_1)}{T_2(t_2) - T_1(t_1)} \\
&= \lim_{t_1 \to t_2} \frac{T_1(t_1)T_2(t_2) - \kappa_1(t_1)\gamma_2(t_2)N_1(t_1)}{-\kappa_1(t_1)N_1(t_1)} \\
&= \gamma_2(t_2) + \frac{1}{\kappa_1(t_2)} N_2(t_2).
\end{aligned}$$

This is the centre of the osculating circle at the point $\gamma_2(t_2)$. We can extend the map C_1 in this way and its is still smooth.

The second centre map $C_1 : I \times I \to \mathbb{R}^2 \cong \mathbb{C}$ is defined by

$$C_2(t_1, t_2) = \frac{\gamma_1(t_1)T_2(t_2) + \gamma_2(t_2)T_1(t_1)}{T_1(t_1) + T_2(t_2)}$$

for all t_1, t_2 such that $T_1(t_1) \neq -T_2(t_2)$. About this point there is a rotation $R_2 : \mathbb{R}^2 \to \mathbb{R}^2$ such that $R_2(\gamma_1(t_1)) = \gamma_2(t_2)$ and $R_2'(T_1(t_1)) = -T_2(t_2)$ (fig. 2.3).

Both these rotations send the line tangent to γ_1 at $\gamma_1(t_1)$ onto the line tangent to γ_2 at $\gamma_2(t_2)$ but the two images are rotated about 180°. To get 3-point or higher contact we need to choose the pair (t_1, t_2) such that $R_i(\gamma_1)$ and γ_2 have the same osculating circle at the point $\gamma_2(t_2)$. For the two osculating circles to have the same radius we require that the curvatures are equal in magnitude i.e. $|\kappa_1(t_1)| = |\kappa_2(t_2)|$. We also need to choose the right centre map to map the osculating circles onto each other.

First assume that $\kappa_1(t_1) = +\kappa_2(t_2)$. Let $u_1(t_1)$ be the centre of curvature of γ_1 at $\gamma_1(t_1)$.

$$\begin{aligned} u_1(t_1) &= \gamma_1(t_1) + 1/\kappa_1(t_1)N_1(t_1) \\ &= \gamma_1(t_1) + i/\kappa_1(t_1)T_1(t_1). \end{aligned}$$

Similarly let the centre of curvature of γ_2 at $\gamma_2(t_2)$ be

$$u_2(t_2) = \gamma_2(t_2) + i/\kappa_2(t_2)T_2(t_2).$$

Now

$$\begin{aligned} R_1(u_1(t_1)) &= (\gamma_1(t_1) + i/\kappa_1(t_1)\ T_1(t_1) - C_1(t_1, t_2))\ e^{i\theta} + C_1(t_1, t_2) \\ &= (\gamma_1(t_1) - C_1(t_1, t_2))\ e^{i\theta} + C_1(t_1, t_2) + i/\kappa_1(t_1)T_1(t_1)e^{i\theta} \\ &= \gamma_2(t_2) + i/\kappa_1(t_1\ T_2(t_2) \\ &= u_2(t_2). \end{aligned}$$

So if $\kappa_1(t_1) = \kappa_2(t_2)$ then the rotation about $C_1(t_1, t_2)$ will ensure that the centres of the osculating circles are mapped onto each other and that we have 3-point contact between $R_1(\gamma_1)$ and γ_2. Now if $\kappa_1(t_1) = -\kappa_2(t_2)$ we use the rotation about $C_2(t_1, t_2)$ and find that

$$\begin{aligned} R_2(u_1(t_1)) &= (\gamma_1(t_1) + i/\kappa_1(t_1)\ T_1(t_1) - C_2(t_1, t_2))\ e^{i\theta} + C_2(t_1, t_2) \\ &= (\gamma_1(t_1) - C_2(t_1, t_2))\ e^{i\theta} + C_2(t_1, t_2) + i/\kappa_1(t_1)\ T_1(t_1)\ e^{i\theta} \\ &= \gamma_2(t_2) - i/\kappa_1(t_1)\ T_2(t_2) \\ &= u_2(t_2), \end{aligned}$$

so we have 3-point contact. We can now define the rotational symmetry set.

Definition 2.2.2

The **Rotational Symmetry Set** (RSS) of two curves γ_1, γ_2 consists of two parts

$$Part\ 1 = \{C_1(t_1,t_2) : \forall\ t_1,t_2 \text{ such that } \kappa_1(t_1) = \kappa_2(t_2)\}$$

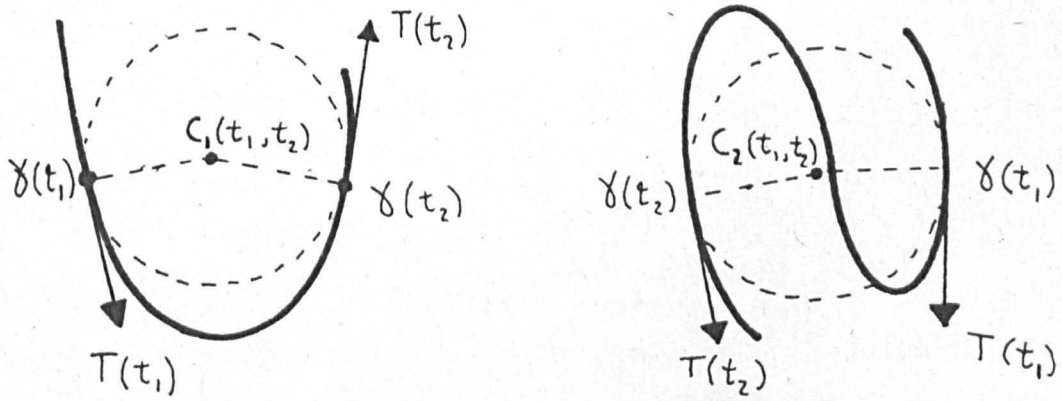
$$Part\ 2 = \{C_2(t_1,t_2) : \forall\ t_1,t_2 \text{ such that } \kappa_1(t_1) = -\kappa_2(t_2)\}$$

This definition also applies when $\gamma_1 = \gamma_2$ and we have the rotational symmetry set of a single curve in which case we add the limit points as $t_1 \to t_2$. In general the Rotational Symmetry Set consists of the union of a number of singular curves and we can predict where the end points, inflections and singularities occur. We will defer a statement about these until after we have defined the symmetry set as the two sets are closely related.

The concept of the RSS (part 1) was thought of by Peter Giblin as an extension of the SS to cover rotational symmetry. The need to study the second part of the set needs was realised jointly by Farid Tari and I.

§2.3 Defining the Symmetry Set and Mid-Point Locus

The symmetry set of a curve $\gamma : I \to \mathbb{R}^2$ is the set of centres of circles which are tangent to γ in two places $\gamma(t_1)$, $\gamma(t_2)$ or have at least 4-point contact with γ somewhere. If such a circle exists then there is a rotation about the centre which sends $\gamma(t_1)$ onto $\gamma(t_2)$. This rotation will also send the line tangent to γ at $\gamma(t_1)$ onto the line tangent to γ at $\gamma(t_2)$ as both lines are tangent to the circle. Hence the



a) Same direction round circle. b) Opposite directions round circle.

fig. 2.4) The orientation of the tangents to the curve with respect to the bi-tangent circle.

centre of the circle is given by one of the centre maps. To decide which we need

to look at the global orientation of the curve (fig. 2.4). We can have one of two situations:

a) Both tangent vectors point in the same direction round the circle (both clockwise or both anticlockwise);

b) The tangent vectors point in opposite directions round the circle (one clockwise and one anticlockwise).

In case a) the centre of the circle is given by $C_1(t_1, t_2)$ the first centre map and in case b) the centre is given by the second centre map $C_2(t_1, t_2)$.

Definition 2.3.1

Let

$$SS_1^{-1} = \left\{ (t_1, t_2) \in I \times I \right. \quad \text{such that there exists a circle tangent to } \gamma \text{ at } \gamma(t_1) \text{ and } \gamma(t_2) \text{ and } T(t_1), T(t_2) \text{ both point in the same direction round the circle.}$$

and

$$SS_2^{-1} = \left\{ (t_1, t_2) \in I \times I \right. \quad \text{such that there exists a circle tangent to } \gamma \text{ at } \gamma(t_1) \text{ and } \gamma(t_2) \text{ and } T(t_1), T(t_2) \text{ point in opposite directions round the circle.}$$

The **Symmetry Set** can be defined in two parts

$$Part\ 1 = SS_1 = \{C_1(t_1, t_2) \mid (t_1, t_2) \in SS_1^{-1}\}$$
$$Part\ 2 = SS_2 = \{C_2(t_1, t_2) \mid (t_1, t_2) \in SS_2^{-1}\}$$

The **Mid Point Locus** is the set

$$\{\tfrac{1}{2}(\gamma(t_1) + \gamma(t_2)) \mid (t_1, t_2) \in SS_1^{-1} \cup SS_2^{-1}\}.$$

The **Evolute** is the set

$$\{C_1(t, t) \mid t \in I\},$$

i.e. the set of centres of circles with at least 3-point contact.

The Symmetry Set and Midpoint Locus have been known for some time [Giblin-Brassett], [Bruce-Giblin-Gibson]. However previous studies have been mainly local where only part 1 of the set is of interest. In this work we look at the set from a more global perspective and we consider both parts of the set separately.

§2.4 The local structure of the Rotational Symmetry Set.

Proposition 2.4.1 [Tari]

We have the following local structure for the first part of the Rotational Symmetry Set of a curve γ.

1) A smooth curve without inflections if

$$\begin{aligned} \kappa(t_1) &= \kappa(t_2), \\ 0 \neq \kappa'(t_1) &\neq \kappa'(t_2) \neq 0, \\ (t_1, t_2) &\notin SS_1^{-1}. \end{aligned}$$

2) A smooth curve with a **Type I** inflection if

$$\begin{aligned} \kappa(t_1) &= \kappa(t_2), \\ \kappa'(t_1) &= 0, \\ \kappa'(t_2) &\neq 0, \\ \kappa''(t_1) &\neq 0, \\ (t_1, t_2) &\notin SS_1^{-1}. \end{aligned}$$

3) A smooth curve with a **Type II** inflection if

$$\begin{aligned} \kappa(t_1) &= \kappa(t_2), \\ (t_1, t_2) &\in SS_1^{-1}, \\ 0 \neq \kappa'(t_1) &\neq \kappa'(t_2) \neq 0. \end{aligned}$$

4) An ordinary cusp if

$$\begin{aligned} \kappa(t_1) &= \kappa(t_2), \\ \kappa'(t_1) &= \kappa'(t_2) \neq 0, \\ (t_1, t_2) &\notin SS_1^{-1}. \end{aligned}$$

5) An end point of a smooth curve if

$$\begin{aligned} t_1 &= t_2, \\ \kappa'(t_1) &= 0, \\ \kappa'' &\neq 0. \end{aligned}$$

i.e. at a vertex of the curve. The end point will lie at the centre of the osculating circle of the vertex and the RSS will be tangent to the normal (fig 2.5).

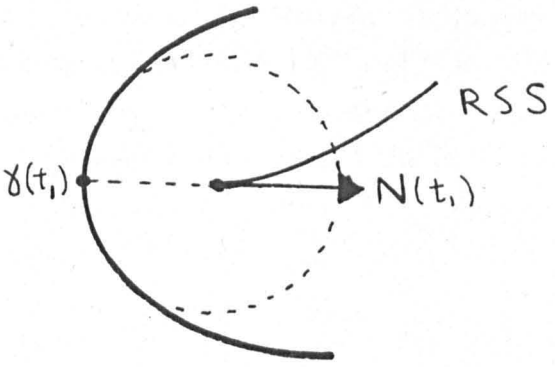
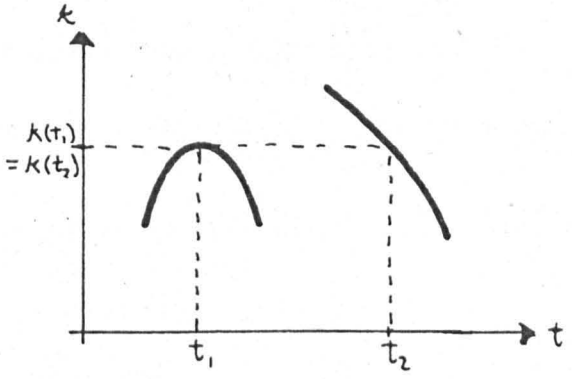
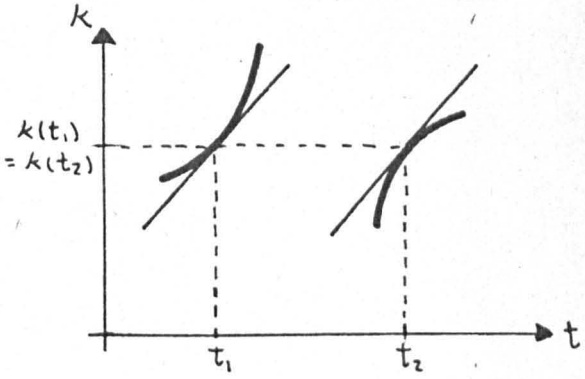


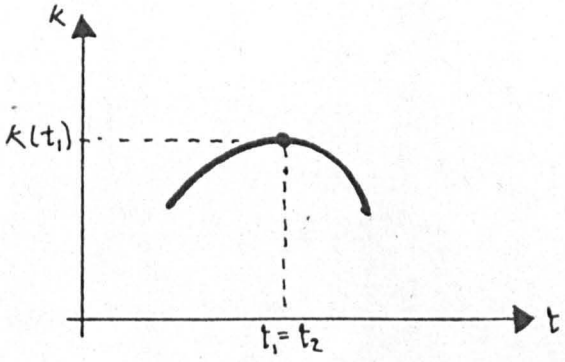
fig. 2.5) An end point on the Rotational Symmetry Set.



Case 2) Type I inflection.



Case 4) Ordinary cusp.



Case 5) End point.

fig. 2.6) Graphs of curvature plotted against arc length for the different cases in prop. 2.4.1).

Cases 2),4) and 5) can be illustrated by looking at the graphs of curvature plotted against arc length (fig 2.6). In case 2) there is an extremum at t_1, in case 4) the gradients at t_1 and t_2 are equal and in case 5) there is a single extremum.

The special points (cases 2-5) on the RSS correspond to additional symmetries of the curve: type II inflections occur when the point lies on both the RSS and

SS hence there is both reflectional and rotational symmetry here; a cusp 4) occurs when there is 4-point contact between $R_1(\gamma)$ and γ at $\gamma(t_2)$ so represents a very strong rotational symmetry; when the derivative of curvature is zero at a point there is 4-point contact between the curve and the osculating circle so end points 5) indicate a very circular part of the curve (vertex); type I inflection indicate rotational symmetry between a vertex of the curve and some other point on the curve which is not a vertex.

Proposition 2.4.2

We have the following local structure for the second part of the Rotational Symmetry Set of a generic curve γ.

1) A smooth curve without inflections if

$$\begin{aligned} \kappa(t_1) &= -\kappa(t_2), \\ 0 \neq \kappa'(t_1) &\neq \kappa'(t_2) \neq 0, \\ (t_1, t_2) &\notin SS_2^{-1}. \end{aligned}$$

2) A smooth curve with a **Type I** inflection

$$\begin{aligned} \kappa(t_1) &= -\kappa(t_2), \\ \kappa'(t_1) &= 0, \\ \kappa'(t_2) &\neq 0, \\ \kappa''(t_1) &\neq 0, \\ (t_1, t_2) &\notin SS_2^{-1}. \end{aligned}$$

3) A smooth curve with a **Type II** inflection

$$\begin{aligned} \kappa(t_1) &= -\kappa(t_2), \\ (t_1, t_2) &\in SS_1^{-1}, \\ 0 \neq \kappa'(t_1) &\neq \kappa'(t_2) \neq 0, \end{aligned}$$

4) An ordinary cusp if

$$\begin{aligned} \kappa(t_1) &= -\kappa(t_2), \\ \kappa'(t_1) &= \kappa'(t_2) \neq 0, \\ (t_1, t_2) &\notin SS_2^{-1}. \end{aligned}$$

5) An end point of a smooth curve if

$$t_1 = t_2,$$
$$\kappa(t_1) = 0,$$
$$\kappa'(t_1) \neq 0.$$

This end point will lie at the inflection of the curve and the RSS will be tangent to the curve at this point (fig 2.7).

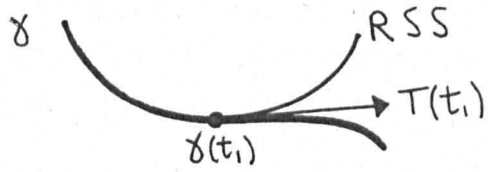


fig. 2.7) An end point of part 2 of the Rotational Symmetry Set.

Proof

The above results are very similar to those for the RSS part I and we can derive results 1) - 4) from the previous proposition.

First divide γ up into two pieces of curve γ_1, γ_2 such that near t_1 we have

$$\gamma_1(\delta t_1) = \gamma(t_1 + \delta t_1)$$

and near t_2

$$\gamma_2(\delta t_2) = \gamma(t_2 + \delta t_2).$$

Now reverse the orientation of one of the curves i.e. define a curve

$$\overline{\gamma_2}(\delta t_2) = \gamma_2(-\delta t_2) = \gamma(t_2 - \delta t_2).$$

Let C_1, C_2 be the two centre maps with respect to the curves γ_1, γ_2 and let $\overline{C_1}, \overline{C_2}$ be the two centre maps with respect to $\gamma_1, \overline{\gamma_2}$. Let $\overline{\kappa_2}$ be the curvature and $\overline{T_2}$ the unit tangent vector for $\overline{\gamma_2}$. We have

$$\overline{\kappa_2}(t_2) = -\kappa_2(-t_2)$$
$$\overline{\kappa_2'}(t_2) = +\kappa_2'(-t_2)$$
$$\overline{T_2}(t_2) = -T_2(-t_2)$$
$$\overline{C_1}(t_1, t_2) = C_2(t_1, -t_2)$$
$$\overline{C_2}(t_1, t_2) = C_1(t_1, -t_2).$$

Now

"The second part of the RSS for curves γ_1, γ_2"

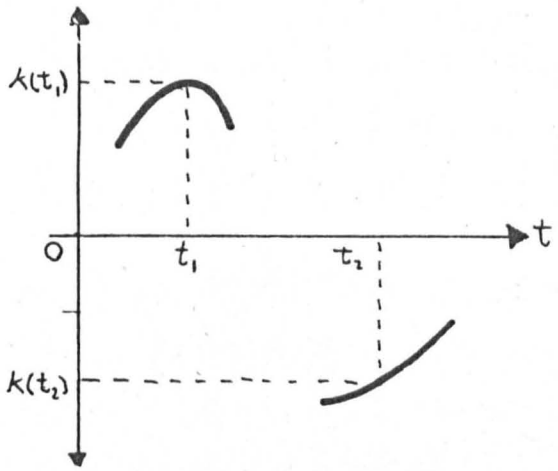
$$= \{C_2(t_1, t_2) | \kappa_1(t_1) = -\kappa_2(t_2)\}$$

$$= \{\overline{C_1}(t_1, -t_2) | \kappa_1(t_1) = \overline{\kappa_2}(-t_2)\}$$

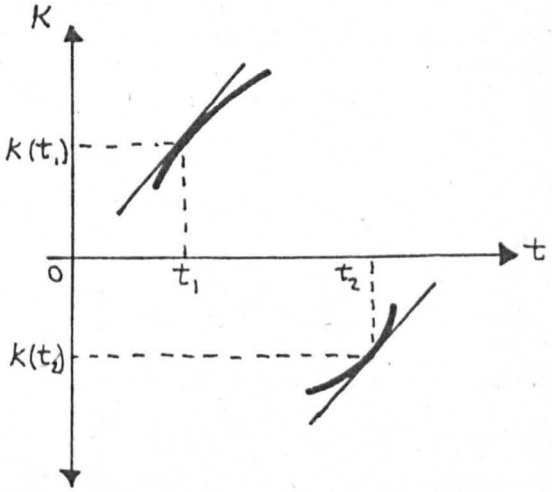
= "first part of the RSS for curves $\gamma_1, \overline{\gamma_2}$".

Hence to deduce a result about the structure of the second part of the RSS we can use results from the first part where the orientation of one piece has been reversed.

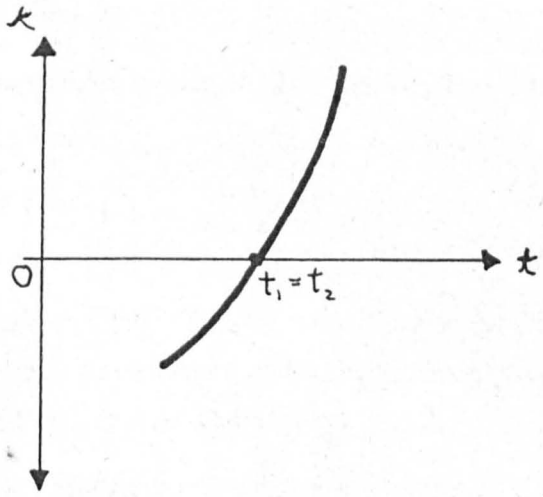
This technique will not work in case 5) where $t_1 = t_2$ and we can not divide the curve into two pieces. The computer pictures shown later gave the first clue as to the structure of the RSS here, these helped in the mathematical proof [Tari].



Case 2) Type I inflection.



Case 4) Ordinary cusp.



Case 5) End point.

fig. 2.8) Graphs of curvature plotted against arc length for the different cases in prop. 2.4.2).

The above conditions can again be illustrated on the graphs of curvature plotted against arc length (fig 2.8). Cases 2) and 4) are very much like those for proposition 2.4.1 except the curvature of one piece is opposite in sign. In 4) the gradients are still equal. Case 5) is just an inflection on the curve. Taking the whole of the RSS all the special points on the curve (vertices and inflections) are indicated by end points on the RSS

§2.5 The local structure of the Symmetry Set and Mid-Point Locus

Proposition 2.5.1

For a generic curve γ we have the following local structure for the Symmetry Set.

1) A smooth curve without inflections if

$$(t_1, t_2) \in SS_1^{-1},\ \kappa(t_1) \neq \kappa(t_2)$$
$$\text{or } (t_1, t_2) \in SS_2^{-1},\ \kappa(t_1) \neq -\kappa(t_2),$$

and if the bitangent circle does not osculate at either point.

2) An inflection if

$$(t_1, t_2) \in SS_1^{-1},\ \kappa(t_1) = \kappa(t_2),$$
$$\text{or } (t_1, t_2) \in SS_2^{-1},\ \kappa(t_1) = -\kappa(t_2),$$

the bitangent circle does not osculate at either point and $\kappa'(t_1) \neq -\kappa'(t_2)$ which prevents higher inflections.

3) An ordinary cusp in the A_1A_2 case i.e. the bitangent circle osculate at $\gamma(t_1)$ but not at $\gamma(t_2)$.

4) An end point if

$$t_1 = t_2 \quad \text{and} \quad \kappa'(t_1) = 0$$

i.e. at a vertex of γ. This can only occur on part 1 of the SS. The end-point of the SS lies at the centre of the osculating circle and the SS is tangent to the normal of the curve at the vertex.

5) A triple crossing when there is a tri-tangent circle.

For proof of these results see [Giblin-Brassett] and [Tari]. We have added explicit conditions for the second part of the set.

We observe that condition 1) is the same as for a type II inflection on the RSS. In fact both these inflections lie at the same point and they are tangent to one another. The end points of part 1 of the RSS and those of the SS also occur in the same places as each other, at the centres of the osculating circles of the vertices. Here they will have the same tangent lines which is the normal to the curve at the vertex (fig. 2.9).

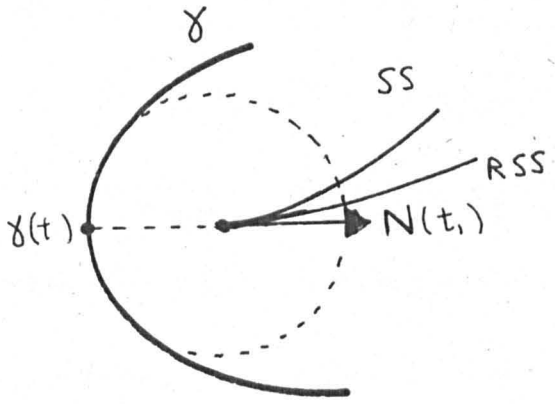


fig. 2.9) An end point on the Symmetry Set.

Proposition 2.5.2

For a generic curve the Mid-Point locus is smooth and has an end point corresponding to each vertex of the curve which lies at the vertex. For more detail see [Giblin-Brassett].

§2.6 An Example of the various sets.

Most of the features of the local structure of the three sets are illustrated by the following computer generated examples, figure 2.10 shows the Rotational Symmetry Set, Symmetry Set and evolute of a curve, figure 2.11 shows the Mid-Point Locus of the same curve and figure 2.12 shows an enlargement of fig 2.10.

In these and the following computer generated examples we use the following colour coding.

2.6.1 Key to Examples

Original Curve	Black
Rotational Symmetry Set	Red or black
Symmetry Set	Blue
Midpoint Locus	Red

Evolute Green

Special points on original curve are marked as follows

Vertices

Inflections

On the Rotational Symmetry Set

Type I inflections

Type II inflections

On the Symmetry Set

Inflections

In this example we note the following features:

a) the RSS (part 1) and the SS have end points in the cusps of the evolute which lie at the centers of the osculating circles of the vertices, the MPL has end points at the vertices of the curve;

b) the RSS (part 2) has a small section between the two inflections with end point at the inflections;

c) There are cusps on both the RSS and SS but the MPL is smooth.

d) there are several places where both the SS and RSS have inflections and at these points the two curves have the same tangent line, the RSS also has type I inflections;

e) the RSS has a point at infinity which corresponds to two points having equal curvatures and equal tangents;

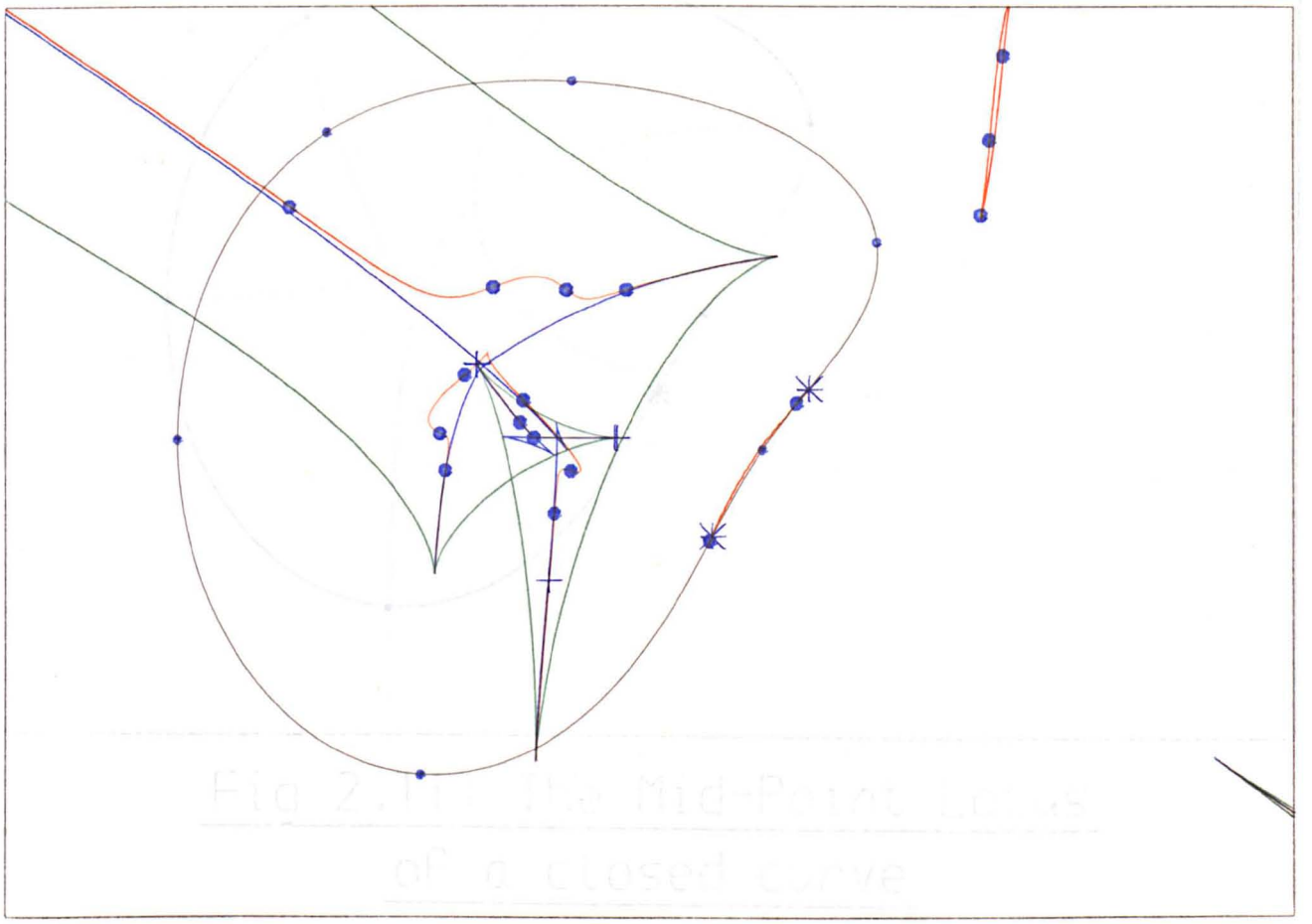
f) the SS has a point a infinity which corresponds or to a bitangent line;

g) the Midpoint Locus of a bounded curve never has points at infinity;

g) there is closed loop on the right of fig. 2.10) which is in part 2 of the RSS. This illustrates quite well the rotational symmetry between the part of the curve between the two inflection and the part near the vertex at the top of the curve;

h) a triple crossing can be seen on the symmetry set.

We see that the sets have a very rich structure which contain alot of information about the curve in the following chapters we will describe the algorithms which enabled such pictures to be drawn.



Key

Curve	—	Vertex of curve	•
RSS	—	Inflection on curve	*
SS	—	Type I inflection on RSS	●
MPL	—	Type II inflection on RSS	+
Evolute	—	= inflection on SS	

Fig 2.10: The Rotational Symmetry Set, and Symmetry Set of a closed curve.

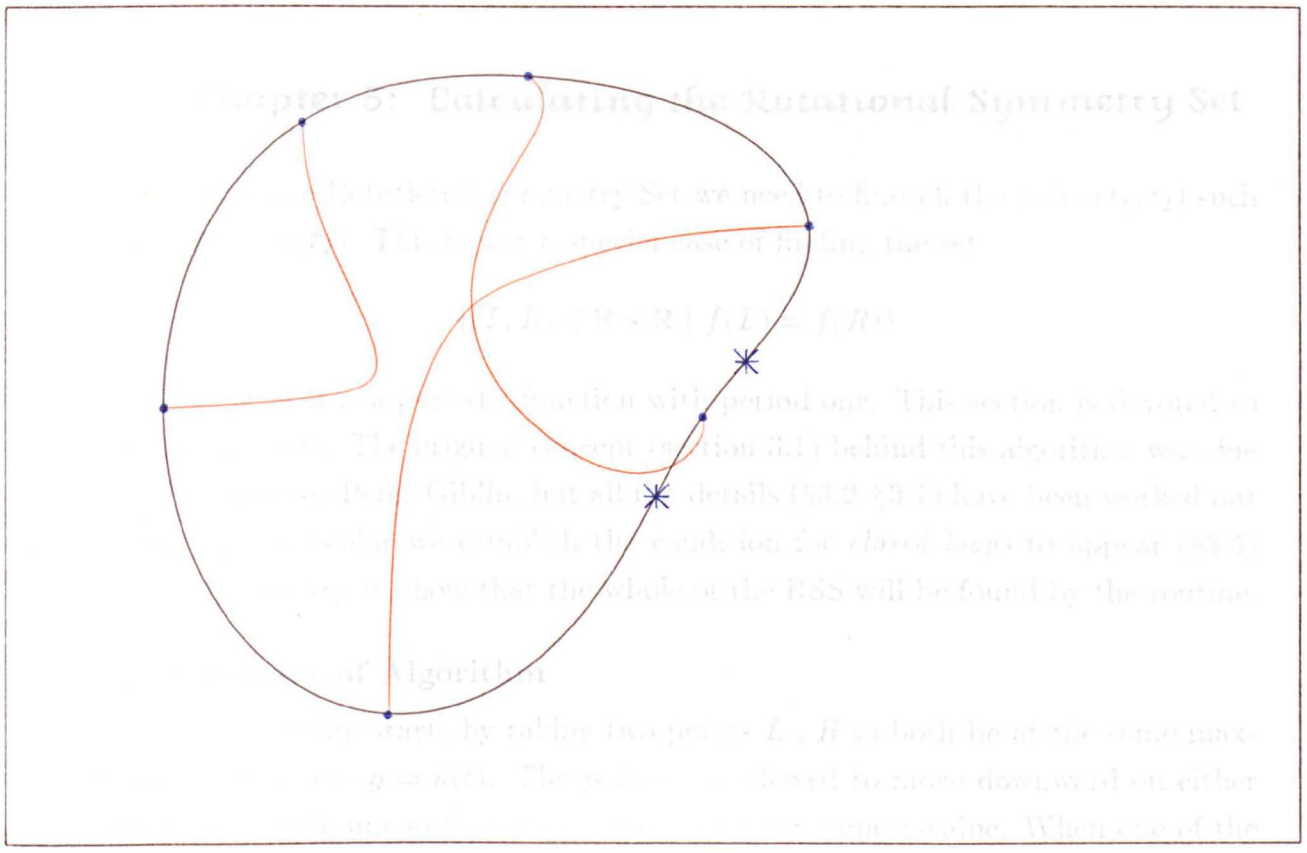


Fig 2.11: The Mid-Point Locus of a closed curve

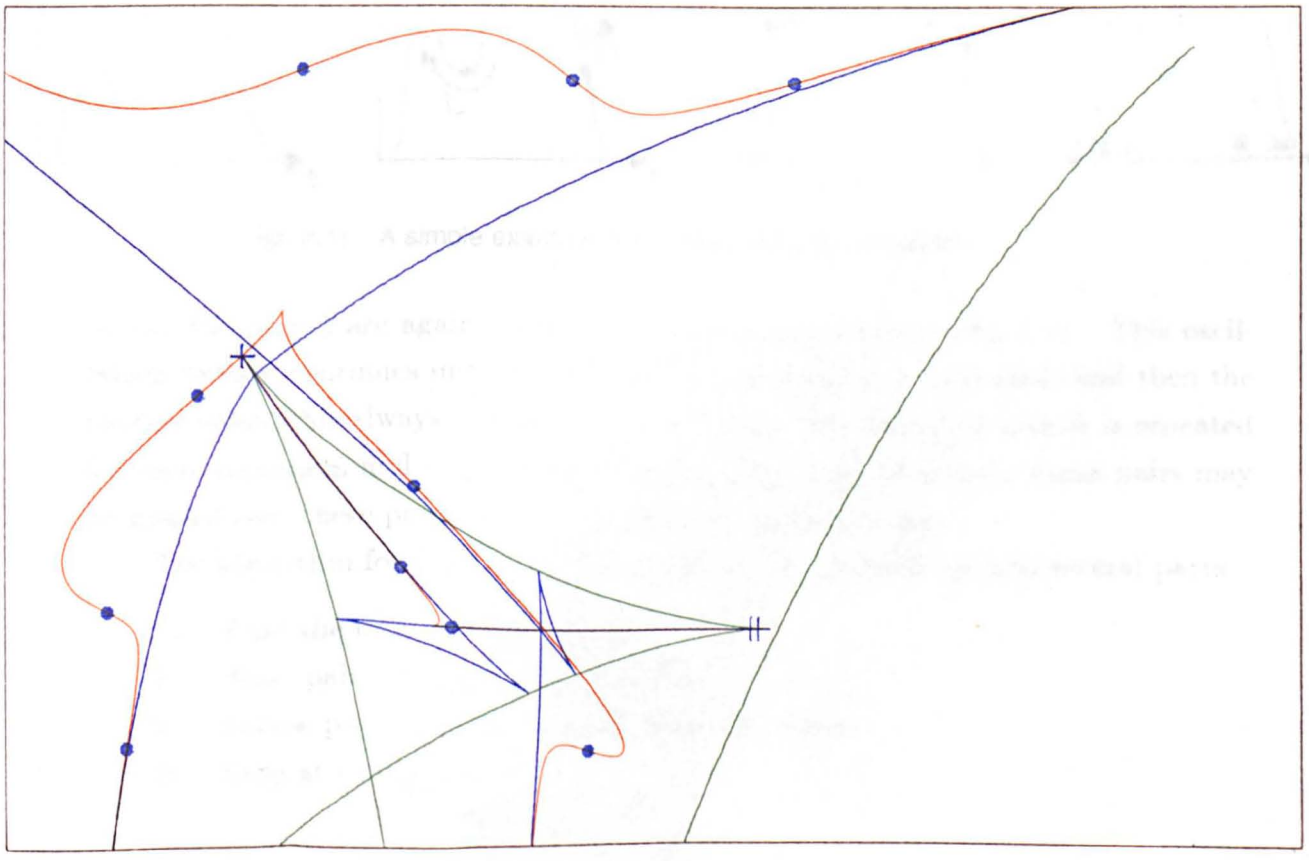


Fig 2.12: An enlargment of Fig 2.10

Chapter 3: Calculating the Rotational Symmetry Set

To find the Rotational Symmetry Set we need to find all the pairs (t_1, t_2) such that $\kappa(t_1) = \kappa(t_2)$. This is just a special case of finding the set

$$\{(L, R) \in \mathbb{R} \times \mathbb{R} \mid f(L) = f(R)\}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a periodic function with period one. This section is devoted to finding such sets. The original concept (section 3.1) behind this algorithm was due to my supervisor Peter Giblin, but all the details (§3.2–§3.7) have been worked out by me. In particular we establish the condition for *closed loops* to appear (§3.5) which is necessary to show that the whole of the RSS will be found by the routine.

§3.1 Summary of Algorithm

The algorithm starts by taking two points L , R to both be at the same maximum of the curve $y = \kappa(t)$. The points are allowed to move downward on either side of the maximum so that they always have the same y-value. When one of the points reaches a minimum the pairs are chosen so that they both have increasing y-values. One of the points will eventually arrive at a maximum and when this

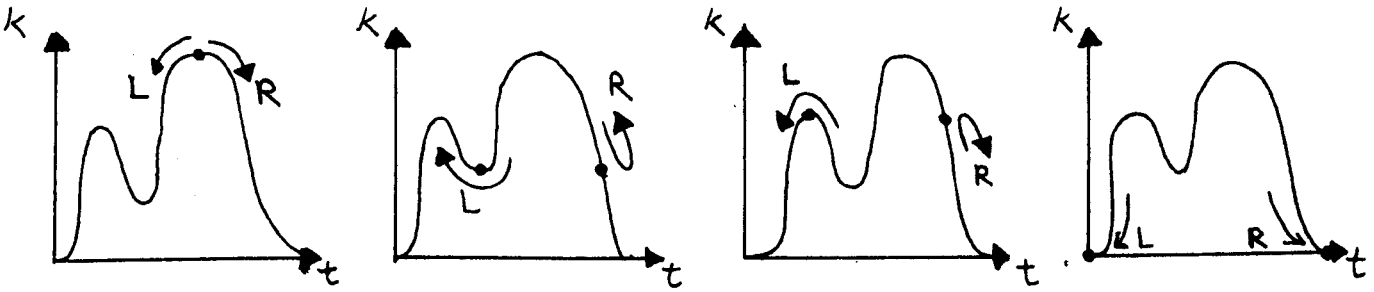


fig. 3.1) A simple example of the movement of the points.

occurs the points are again chosen with decreasing y-values (fig 3.1). This oscillating motion continues until both points are identical (up to period) and then the routine stops, this always happens at a minimum. The above procedure is repeated for each maximum and will eventually cover most pairs of points. Some pairs may be missed out, these pairs can be found by a special routine.

The algorithm for finding all the points can be divided up into several parts

1) Find the maxima and minima.
2) Start pairs of points from maxima.
3) Follow pairs of point up and down the curve.
4) Stop at minima.

5) Do the "Closed Loop" case:- see section 3.5.

The first of these is simply a routine which finds those points for which $\kappa'(t) = 0$ and we will not go into detail about it, the other points are dealt with below.

§3.2 The movement of the points

The movement of the points can be divided up into four subroutines (fig. 3.2)

a) "Descend" = Down and Outwards,
b) "Valley" = Down and Inwards,
c) "Ascend Left" = Up and to the left,
d) "Ascend Right" = Up and to the right.

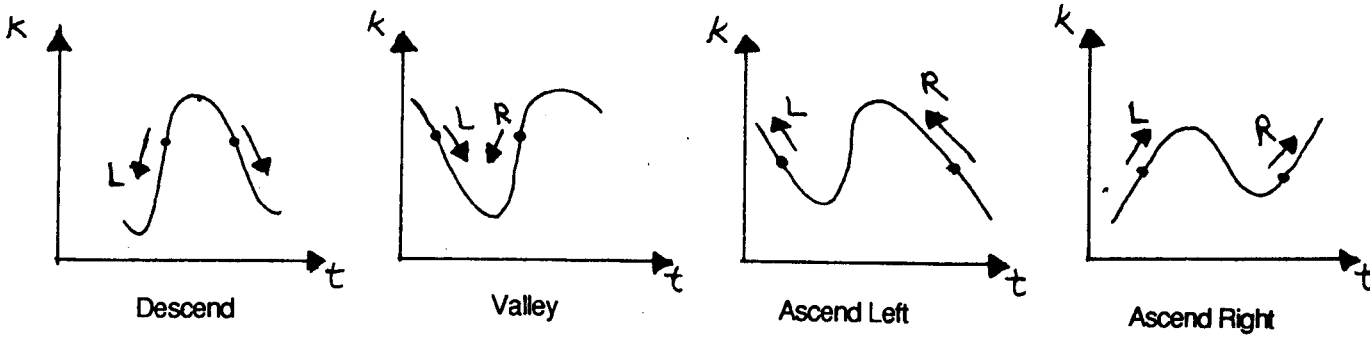


fig. 3.2) The four different types of movement.

Without loss of generality we assume the lowest minimum (of a generic function) in the range $[0, 1)$ lies at $t = 0$. We can then label the two points L and R where L is the left most point (has the lowest t-value). As $t = 0$ is the lowest minimum both points are constrained to lie in the region $[0, 1]$. When they both reach this minimum the routine will stop. Hence there is no ambiguity in the labeling system.

Now we define our terms

Outwards:	point L moves to the left (decreasing value of t)
	point R moves to the right (increasing value of t),
Inwards:	point L moves to the right, R moves to the left,
Leftwards:	Both L,R move left,
Rightwards:	Both L,R move right,
Up:	L,R have increasing y-values,
Down:	L,R have decreasing y-values.

In each of the four movement routines successive pairs of points with equal

y-values are found. Starting with a pair L_0, R_0 which have equal y-values the task of each routine is to find another pair L_1, R_1 close to L_0, R_0 with equal y-values. It was considered to be desirable for the change in the t-values of both points to be kept small. To achieve this we choose a maximum increment allowed "*tinc*" and consider the points $L_0 \pm tinc$, $R_0 \pm tinc$ with signs chosen to fit with the type of movement. If the difference in y-values between L_0 and $L_0 \pm tinc$ is larger than the difference in y-values between R_0 and $R_0 \pm tinc$ i.e.

$$|\kappa(L_0 \pm tinc) - \kappa(L_0)| > |\kappa(R_0 \pm tinc) - \kappa(R_0)| \qquad 3.2.1)$$

then we let $R_1 = R_0 \pm tinc$ and find the point L_1 between L_0 and $L_0 \pm tinc$ with

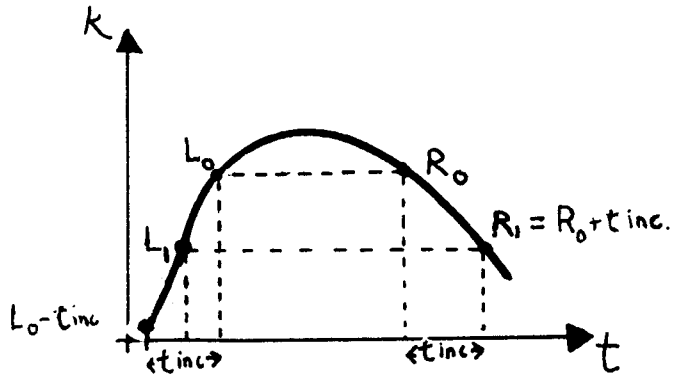


fig. 3.3) Finding the next pair of points in the routine "descend."

the same y-value as R_1 (fig. 3.3). This point can be found by repeated bisection of the interval. If the inequality 3.2.1) is reversed we let $L_1 = L_0 \pm tinc$ and find a point R_1 between $R_0 \pm tinc$ and R_0 with the same y-value as L_1. The above procedure is repeated until one of the two points L, R passes an extremum. This method was chosen in preference to finding points for which the y-values were some fixed increment away from that of L_0: for a function with a shallow slope this would lead to a large change in t-values. Another alternative would have been to fix the change in t-value of one of the points and then find the other point with the same y-value: again this might lead to a large change in the t-value.

For a generic function no two maxima (or two minima) will have the same y-value so there is only one way for the points to keep moving continuously when one arrives at an extremum. This is for the point at the extremum to keep moving in the same t-direction (left or right) and for the other point to reverse its t-direction.

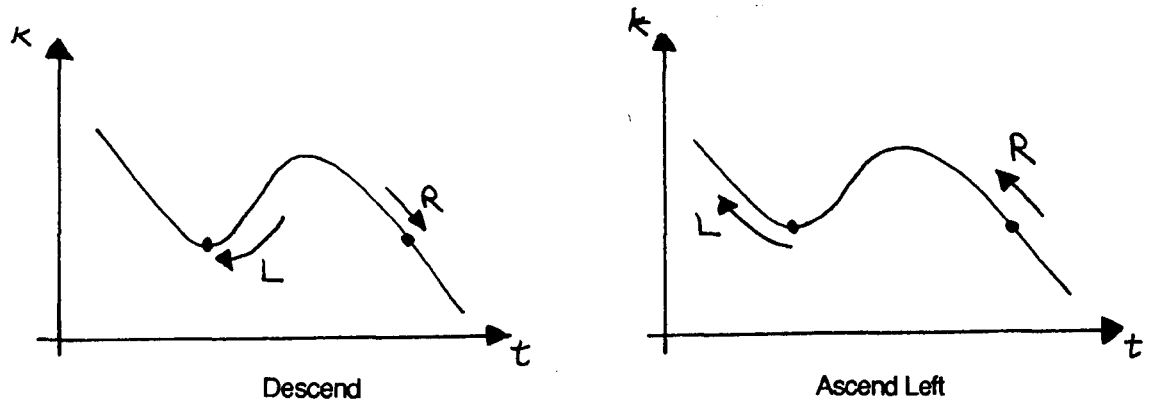


fig. 3.4) The change in type of movement when a minima is reached.

Both points will reverse their y-direction of movement. Figure 3.4 illustrates the change of movement when the left point reaches a minima, the left point continues moving to the left and the right point changes t-direction and move to the left as well. Figure 3.5 shows how we change between the different movement routines are

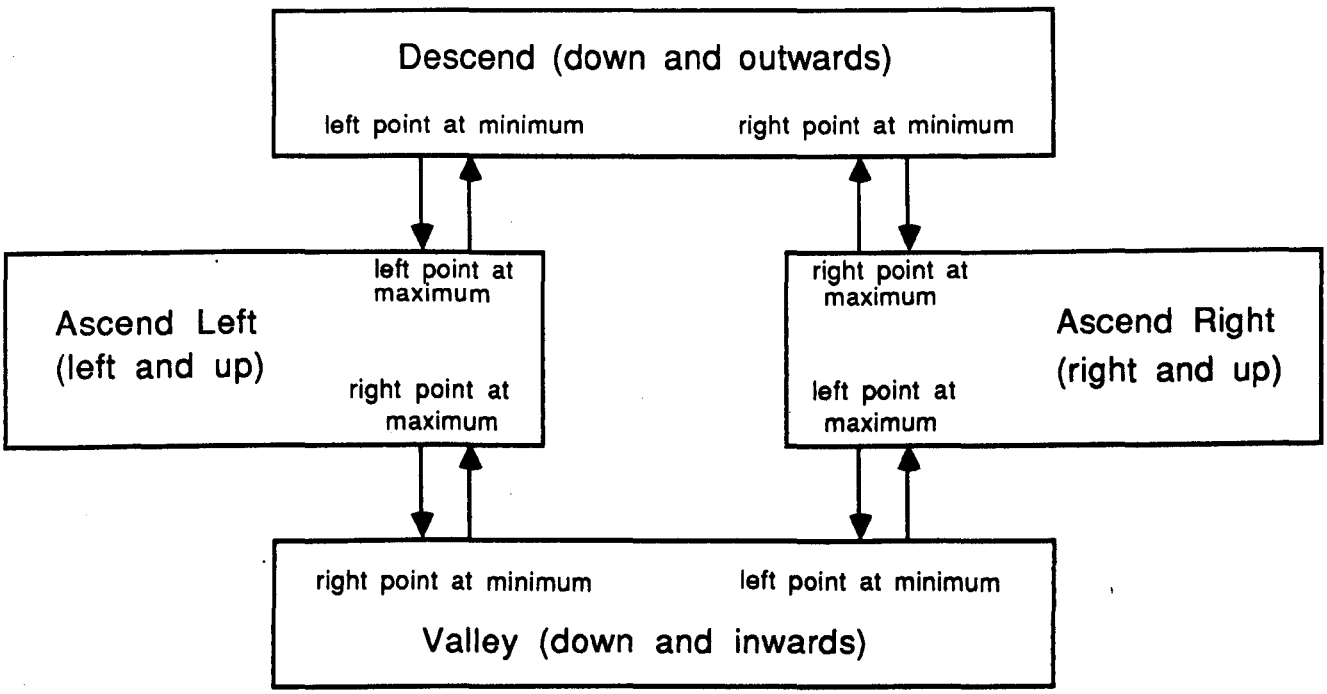


fig 3.5 How the different movement routines are connected

connected. As all possible cases for a generic curve are covered it follows that these four routines are sufficient to cover all possible forms of movement. If the routine had started at a minimum and worked upwards than the four routines would be different i.e.

a) Up and Outwards,
b) Up and Inwards,
c) Down and leftwards,
d) Down and to the right.

The routine will always start at each maximum, M, in turn with $L_0 = R_0 = M$ in the routine "Descend" and can only end when the two points are moving in opposite direction ("Descend" or "Valley"). This will happen when both points reach the same minimum (up to period) from opposite directions. We will see later that when starting at a maximum the pairs will eventually reach the same minimum.

§3.3 Dealing with two extrema of equal y-value.

In a generic family of curves it is possible for two maxima (or two minima) of

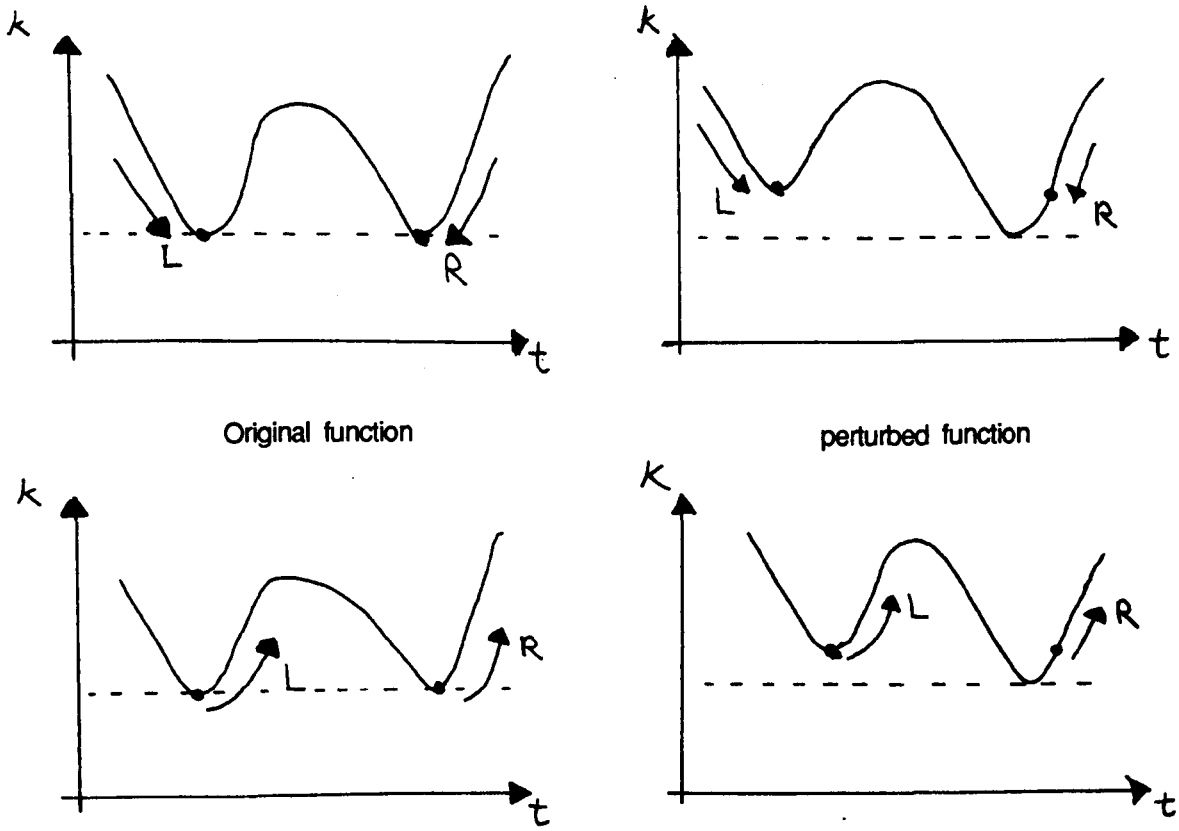


fig. 3.6) The effect of a slight perturbation of the function on the movement of the points.

one of the curves in the family, to have the same y-value. This may also occur in a computer implementation due to the limit of accuracy. When this occurs their

is no natural way to continue moving and a choice needs to be made. This choice must not lead to any pairs of points being missed out.

One way of making this choice is to consider a small perturbation of the curve so that the left most of the two extrema is slightly higher that the right (fig. 3.6). This enables the four routines above to be used. The actual definition of the function does not need to be changed just the decision statements . In one of the two routines which move downwards the small perturbation would mean that the left hand minimum is reached first so in the routine we assume that if both points reach different minima with the same height then only the left hand minimum has been reached. In the other two (upward moving) routines the situation is reversed and if two maxima are reached simultaneously we assume that only the right hand point has reached a maximum. If the routine will cover all pairs of points on the

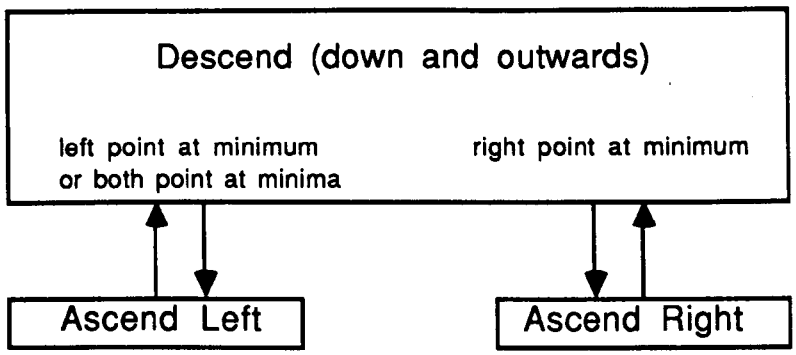


fig 3.7) The revised decision statement for "Descend"

slightly perturbed curve we see that it will cover all pairs on the original curve. The complete decision statement for the routine "Descend" is now shown in figure 3.7.

§3.4 A different way of regarding the problem

A very useful way of thinking about the set of pair (L, R) for which $\kappa(L) = \kappa(R)$ is to define a function $K(t_1, t_2) = \kappa(t_1) - \kappa(t_2)$ and draw the set $K^{-1}(0)$ in the plane (fig. 3.8). As κ is a periodic function opposite edges of the unit square are identified also as K is anti-symmetric , $K(t_1, t_2) = -K(t_2, t_1)$, each un-ordered pair (L, R) is represented twice on this diagram. The line $\{ (t_1, t_2) \mid t_1 = t_2 \}$ is part of the $K^{-1}(0)$ contour and there is a one to one correspondence between points on this line and the curve $y = \kappa(t)$. Furthermore the points $a, \ldots, d$ of figure 3.8 correspond to the extrema of κ and part on the contour always starts at each such point.

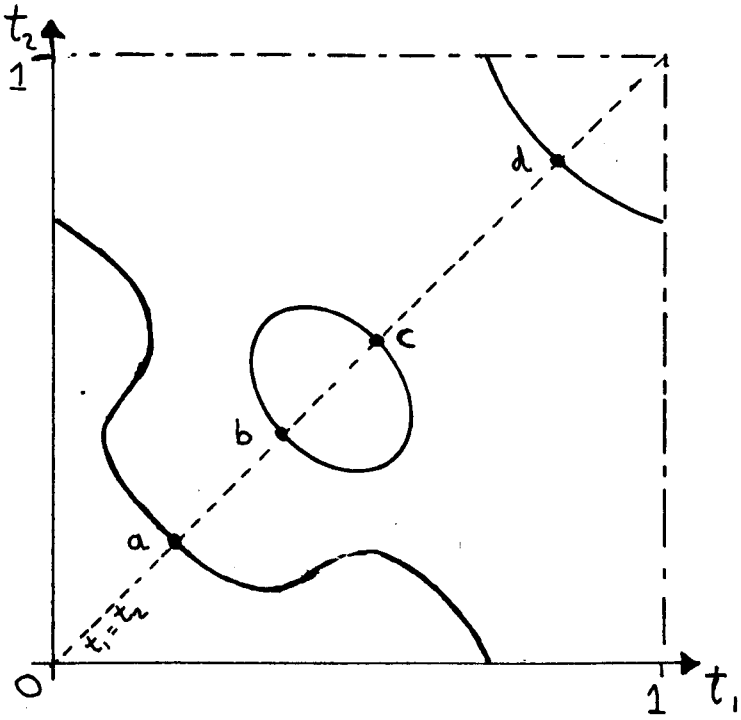


fig. 3.8) The set $K^{-1}(0)$ for a typical curve.

Proposition 3.4.1

The local structure of the $K^{-1}(0)$ contour in each of the following cases is illustrated in figure 3.9.

1) $\kappa'(t_1) > 0$ and $\kappa'(t_2) > 0$.
2) $\kappa'(t_1) < 0$ and $\kappa'(t_2) < 0$.
3) $\kappa'(t_1) > 0$ and $\kappa'(t_2) < 0$.
4) $\kappa'(t_1) < 0$ and $\kappa'(t_2) > 0$.
5) $\kappa'(t_1) = 0$, $\kappa''(t_1) > 0$ and $\kappa'(t_2) > 0$ (min of κ at t_1).
6) $\kappa'(t_1) = 0$, $\kappa''(t_1) < 0$ and $\kappa'(t_2) < 0$ (max of κ at t_1).
7) $\kappa'(t_1) = 0$, $\kappa''(t_1) > 0$ and $\kappa'(t_2) < 0$ (min of κ at t_1).
8) $\kappa'(t_1) = 0$, $\kappa''(t_1) < 0$ and $\kappa'(t_2) > 0$ (max of κ at t_1).
9) $\kappa'(t_2) = 0$, $\kappa''(t_2) > 0$ and $\kappa'(t_1) > 0$ (min of κ at t_2).
10) $\kappa'(t_2) = 0$, $\kappa''(t_2) < 0$ and $\kappa'(t_1) < 0$ (max of κ at t_2).
11) $\kappa'(t_2) = 0$, $\kappa''(t_2) > 0$ and $\kappa'(t_1) < 0$ (min of κ at t_2).
12) $\kappa'(t_2) = 0$, $\kappa''(t_2) < 0$ and $\kappa'(t_1) > 0$ (max of κ at t_2).
13) $t_1 = t_2$, $\kappa'(t_1) > 0$.
14) $t_1 = t_2$, $\kappa'(t_1) < 0$.

15) $t_1 = t_2$, $\kappa'(t_1) = 0$ and $\kappa''(t_1) > 0$ (min of κ at t_1).

16) $t_1 = t_2$, $\kappa'(t_1) = 0$ and $\kappa''(t_1) < 0$ (max of κ at t_1).

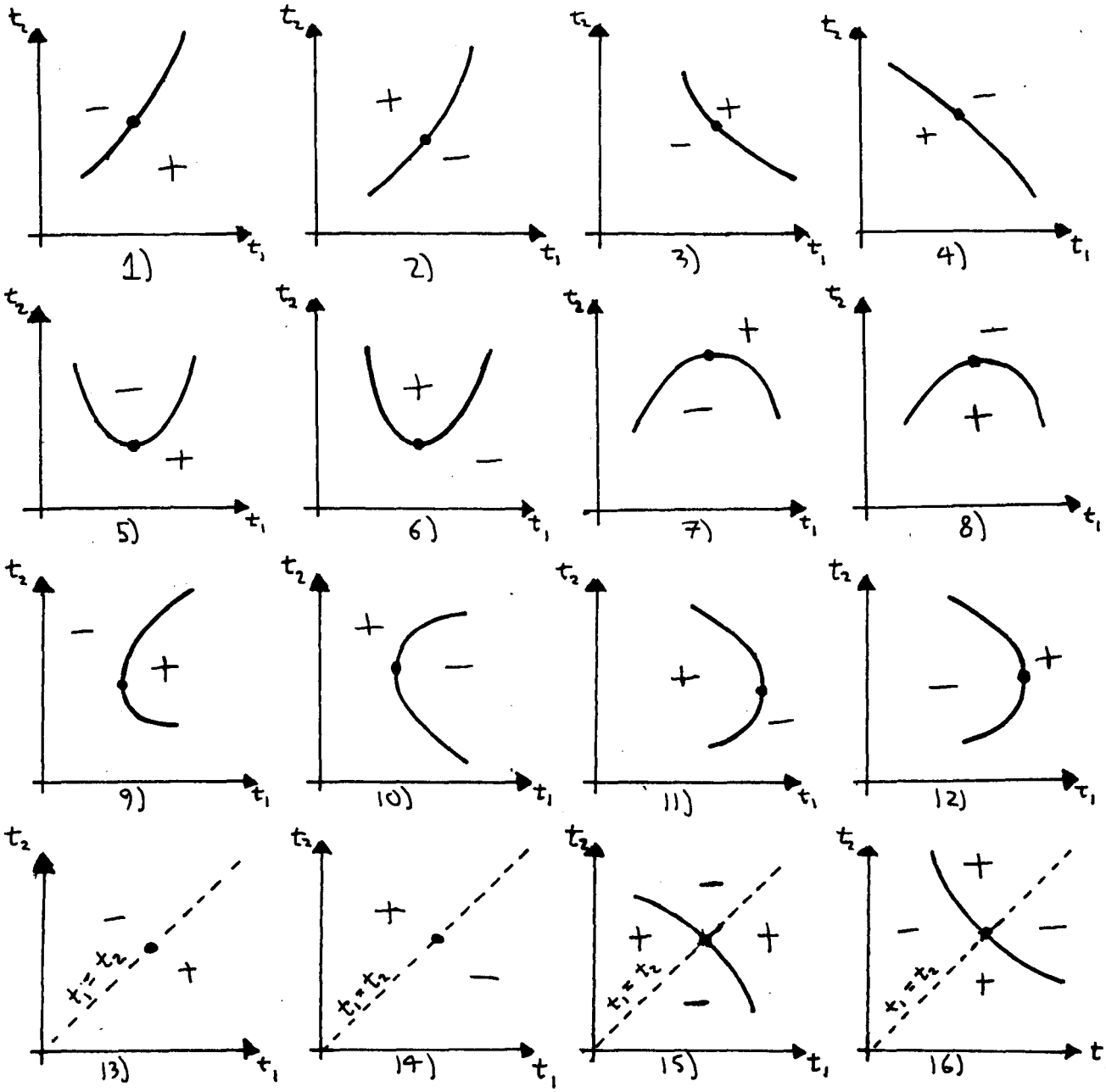


fig. 3.9) The local structure of the $K^{-1}(0)$ contour.

Proof

In each case we approximate $\kappa(t_1)$, $\kappa(t_2)$ near a point (x_0, y_0) on the $K^{-1}(0)$ contour by Taylor polynomials. For cases 1 to 4 we have

$$\kappa(x_0 + x) = \kappa(x_0) + \kappa'(x_0)x + O(x^2)$$
$$\kappa(y_0 + y) = \kappa(y_0) + \kappa'(y_0)y + O(x^2).$$

The set $K^{-1}(0)$ is given by the points $(x_0 + x, y_0 + y)$ such that

$$\kappa(x_0 + x) = \kappa(y_0 + y)$$
$$\iff \kappa(x_0) + \kappa'(x_0)x + O(x^2) = \kappa(y_0) + \kappa'(y_0)y + O(y^2)$$
$$\iff \kappa'(x_0)x(1 + O(x)) = \kappa'(y_0)y(1 + O(y))$$

So in cases 1 to 4 this set is locally given by

$$y \approx \frac{\kappa'(x_0)\, x}{\kappa'(y_0)},$$

we see that the sign of the first derivative affects the direction of the slope. To distinguish between cases 1 and 2 we just move along the line $x \times I$ and consider the sign of K at a nearby point. The other cases all follow similarly.

§3.5 Have all the pairs (L, R) been found?

The above algorithm does not always find all the pairs $(L, R) \in I \times I$ such that

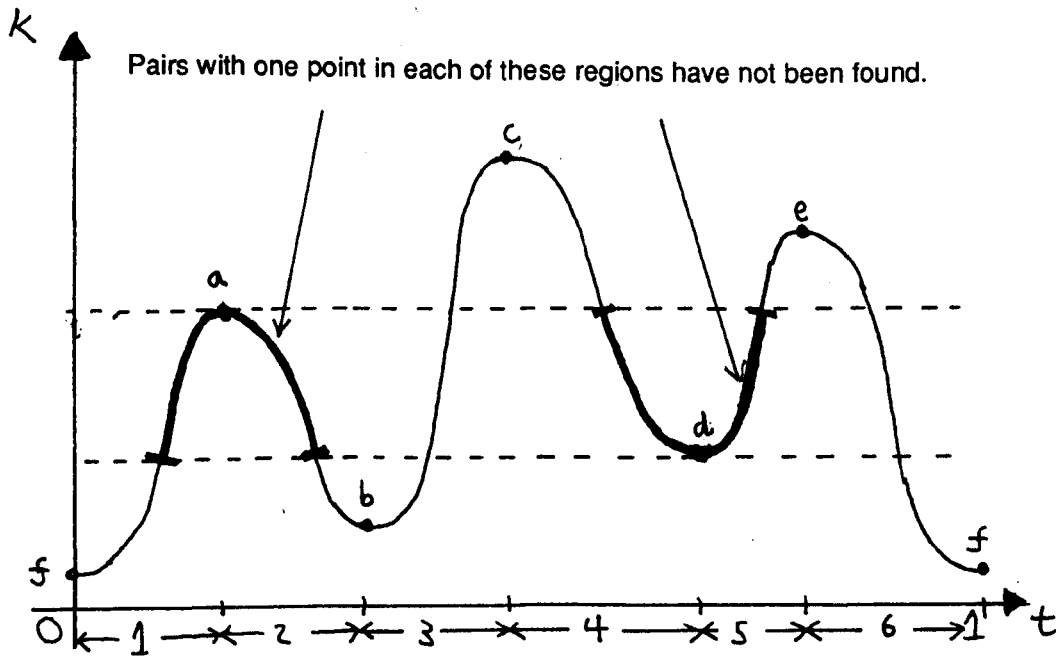


fig. 3.10) A function for which not all pairs have been found.

$\kappa(L) = \kappa(R)$. Consider the function shown in figure 3.10. Starting the routine

with $L = R = a$ we first find all the pairs with L in region 1 and R in region 2 then all pairs in 1 & 3, 2 & 3. Starting at c we find pairs in regions 3 & 4, 3 & 5, 3 & 6, 2 & 6, 1 & 6, and starting at e we find pairs in 5 & 6, 4 & 6, 4 & 5. the routine has omitted all the pairs in regions 1 & 4, 1 & 5, 2 & 5, 2 & 4 which are indicated by the heavy black line in fig 3.10.

To see what has happened we look at the set $K^{-1}(0)$ which is shown in figure

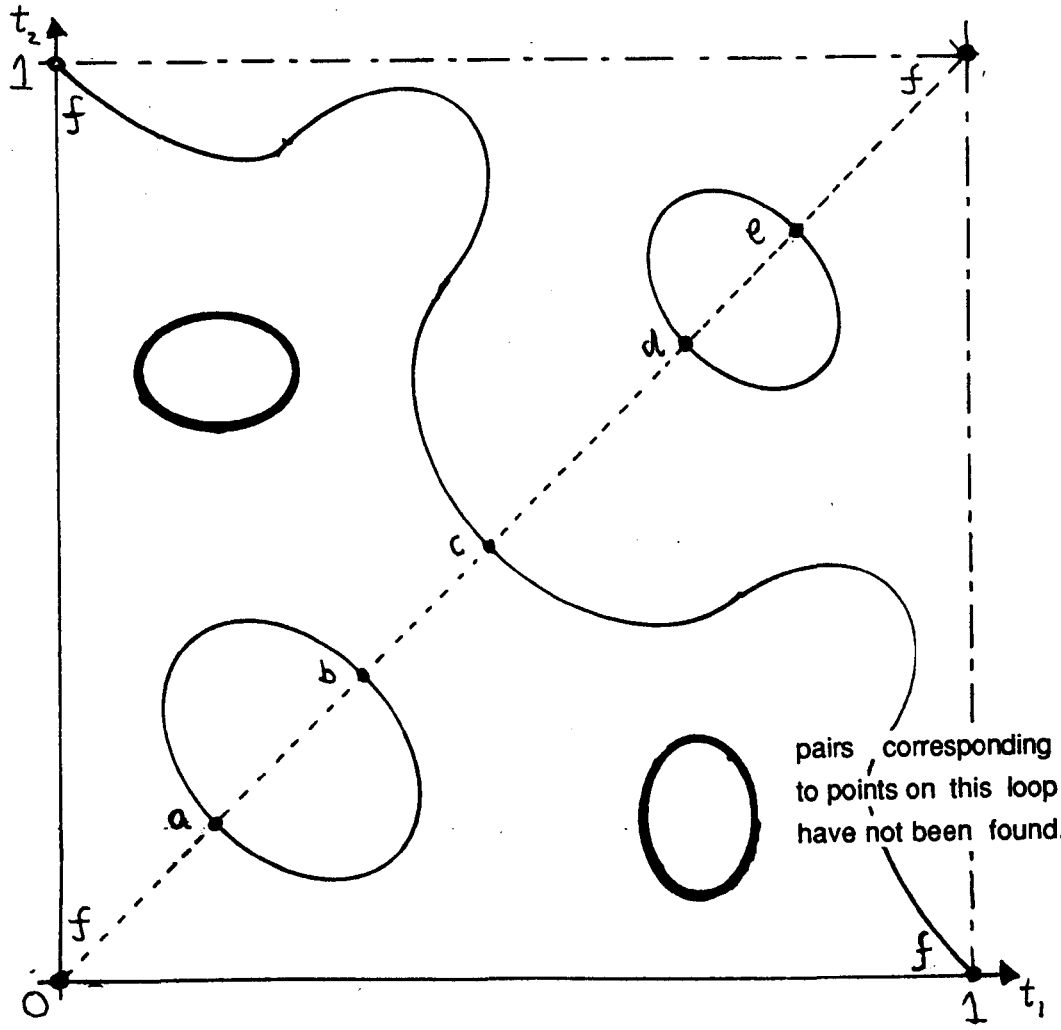


fig. 3.11) The $K^{-1}(0)$ contour for the function shown in fig. 3.10).

3.11. The thin lines corresponds to the part of the $K^{-1}(0)$ contour which can be found using the algorithm as above, these lines start and end at the points $a, \ldots, f$. The heavy circles are also parts of the $K^{-1}(0)$ contour but correspond to pairs of points which have not been found. Note the function K is anti-symmetric so the two heavy circles actually correspond to the same set of pairs of points. To find the missing points we need an algorithm which will find all such isolated loops.

From proposition 3.4.1) we see that for a generic function the $K^{-1}(0)$ contour has no self intersections away from the line $\{t_1 = t_2\}$ so we only have to consider closed loops. If the lowest minimum of κ is at m_0 then the contour can only cross the boundary of the box $[m_0, m_0+1] \times [m_0, m_0+1]$ at the corners which are identified with the point (m_0, m_0). This shows we only have to consider loops contained in a unit square (rather than loops round mobius bands which may occur for the Symmetry Set §4.5). Note as K is anti-symmetric the same loop is represented twice once with a positive interior and once with a negative interior. By finding all the loops with positive interiors we will find all the required points.

Definition 3.5.1

Here we say a **Closed Loop** $\mathcal{L}$ is a component of $K^{-1}(0)$ which does not cross the diagonal $t_1 = t_2$ and has a positive interior. Note by positive interior we are only referring to the sign of points contained inside $\mathcal{L}$ and not also contained in any other *closed loop* nested inside $\mathcal{L}$.

To each loop we associate a box and a point.

Definition 3.5.2

Let

$$M = \{\ a \in I \mid \kappa(t) \text{ has a local maximum at } t = a\ \}$$
$$N = \{\ a \in I \mid \kappa(t) \text{ has a local minimum at } t = a\ \}$$
$$S = \begin{cases} [a_1, a_2] \times [b_1, b_2] \subset I \times I \\ \text{such that } a_1, a_2 \in N,\ b_1, b_2 \in M,. \end{cases}$$

We will call an element of S a **box**. For any *box* $[a_1, a_2] \times [b_1, b_2] \in S$ their is an **associated point** (a_0, b_0) which satisfies the following equations

$$\kappa(a) \leq \kappa(a_0)\ \forall\ a \in [a_1, a_2],$$
$$\kappa(b) \geq \kappa(b_0)\ \forall\ b \in [b_1, b_2].$$

Note for a generic function f the point (a_0, b_0) is uniquely defined for each **box**.

There is a one to one correspondence between the set of *closed loops* and the

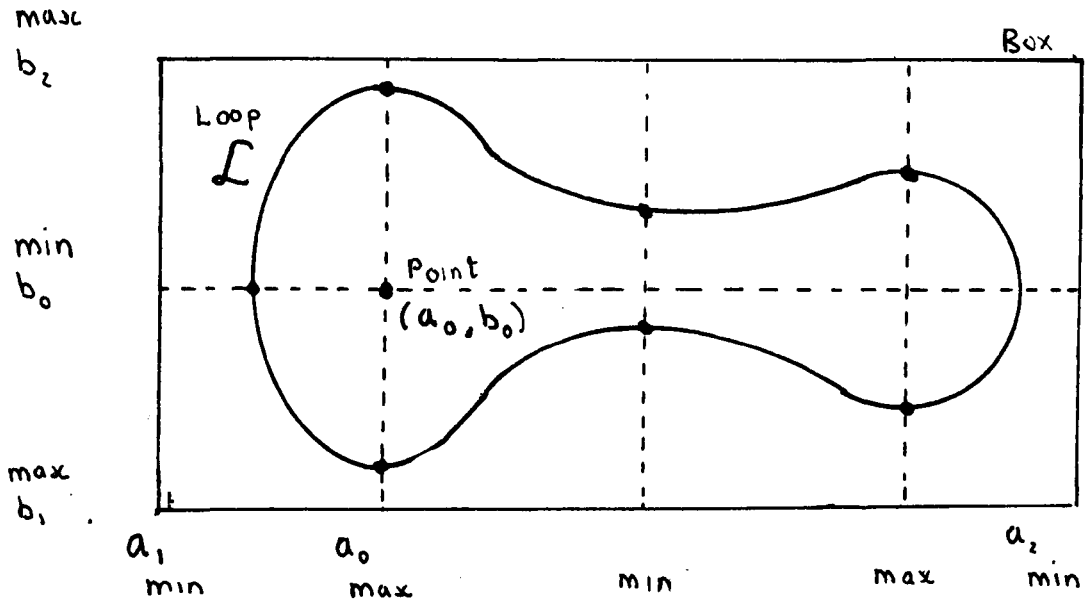


fig. 3.12) A loop and its associated box.

set of boxes and associated points. The following propositions concern the correspondence between *boxes* and *closed loops*.

Proposition 3.5.3

Let $A = [a_1, a_2]$, $B = [b_1, b_2]$ be two intervals such that $A \times B$ is a *box* and let (a_0, b_0) be its *associated point*. If

1) $K(a_0, b_0) > 0$,
2a) $K(a_1, b) < 0$ for all $b \in B$,
2b) $K(a_2, b) < 0$ for all $b \in B$,
2c) $K(a, b_1) < 0$ for all $a \in A$,
2d) $K(a, b_2) < 0$ for all $a \in A$

then there exists a *closed loop* $\mathcal{L} \subset A \times B$.

Proof

This follows immediately form the intermediate value theorem. Note as $K^{-1}(0)$ is non-zero round the perimeter of the *box* neither the *box* or the *loop* can intersect the diagonal.

Proposition 3.5.4

Let $\mathcal{L}$ be a *closed loop*. Choose two intervals $A = [a_1, a_2]$, $B = [b_1, b_2]$ such that $A \times B$ is the smallest *box* in S containing $\mathcal{L}$. Let (a_0, b_0) be the *associated point* of this *box*. We have

1) $K(a_0, b_0) > 0$,
2a) $K(a_1, b) < 0$ for all $b \in B$,
2b) $K(a_2, b) < 0$ for all $b \in B$,
2c) $K(a, b_1) < 0$ for all $a \in A$,
2d) $K(a, b_2) < 0$ for all $a \in A$.

Proof

Let $\alpha_0 = a_1, \alpha_1, \ldots, \alpha_{2m} = a_2$ be the points in $A \cap (M \cup N)$ such that $\alpha_i < \alpha_{i+1}$. As the maxima and minima of a generic function alternate we also have $a_{2i} \in N$ and $a_{2i-1} \in M$. Likewise let $\beta_0 = b_1, \ldots, \beta_{2n} = b_2$ be the points of $B \cap (M \cup N)$ with $\beta_i < \beta_{i+1}$, $\beta_{2i} \in M$ and $\beta_{2i-1} \in N$. As K assumes its absolute maximum value on the *box* at (a_0, b_0) and $\mathcal{L}$ has a positive interior it value must be positive here which proves part 1). We label the extremal points of the *closed loop* as (c_i, d_i) $\{i = 1, \ldots, 4\}$ where (c_1, d_1) is the left most point, (c_2, d_2) the right most point, (c_3, d_3) the point at the bottom of $\mathcal{L}$ and (c_4, d_4) the point at the

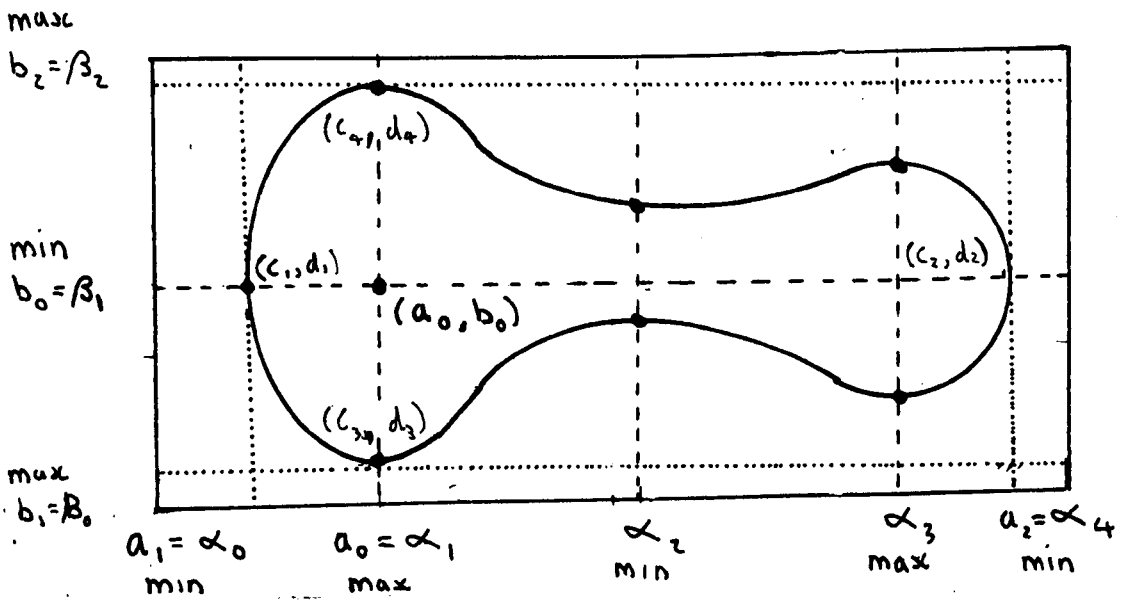


fig 3.13) The points c_i x d_i for the loop in fig 3.12)

top. These points are shown in figure 3.13. By choice of the box $\alpha_0 < c_1 < \alpha_2$ and by proposition 3.4.1) we have $\kappa'(d_1) = 0$, $\kappa'(c_1) > 0$ so $\alpha_0 < c_1 < \alpha_1$ and

$\kappa(c_1) > \kappa(\alpha_0)$. Likewise $\alpha_{2m-1} < c_2 < \alpha_{2m}$, $\beta_0 < d_3 < \beta_1$ and $\beta_{2n-1} < d_4 < \beta_{2n}$. As $d_3 < \beta_1 \leq b_0 \leq \beta_{2n-1} < d_4$ the line $A \times b_0$ must intersect $\mathcal{L}$ at least twice at points (γ_1, b_0), (γ_2, b_0) say. For any crossing point (γ, b_0) we have

$$\begin{aligned} K(\gamma, b) &= \kappa(\gamma) - \kappa(b) \\ &\leq \kappa(\gamma) - \kappa(b_0) \\ &= 0 \end{aligned} \tag{3.5.1}$$

for all $b \in B$ with equality holding only when $b = b_0$. We observe that there can only be two crossing points one on each side of (a_0, b_0) as if (γ_1, b_0), (γ_2, b_0) are two crossing points on the same side of (a_0, b_0) then at least one of the lines $\gamma_1 \times B$ or $\gamma_2 \times B$ must cross $\mathcal{L}$ which contradicts equation 3.5.1) above. We now see that (γ_1, b_0) and (γ_2, b_0) are the extremal points on the loop and we have $c_1 = \gamma_1$, $c_2 = \gamma_2$ and $d_1 = d_2 = b_0$. We repeat the process in the vertical direction where (a_0, δ_3) and (a_0, δ_4) are the two crossing points of $\mathcal{L}$ and $a_0 \times B$ and we have $\delta_3 = d_3$, $\delta_4 = d_4$ and $a_0 = c_3 = c_4$. Along the four lines $c_1 \times B$, $c_2 \times B$, $A \times d_3$ and $A \times d_4$ the function F is negative except for the four points (c_i, d_i) (see fig 3.13). It is now a small step to consider the boundary of the box and we have

$$\begin{aligned} K(a_1, b) &= \kappa(a_1) - \kappa(b) \\ &< \kappa(c_1) - \kappa(b) \\ &\leq \kappa(c_1) - \kappa(b_0) \\ &= 0 \end{aligned}$$

for all $b \in B$ which proves part 2a) the other parts follow the same argument.

For any *closed loop* we can choose the smallest *box* and then we associate a point with this box we call this the **associated point** of the *loop*.

Proposition 3.5.5

No two loops have the same *associated point.*

Proof

Assume that two loops $\mathcal{L}$, $\mathcal{L}'$ have the same associated point (a_0, b_0). Let $A \times B$ be the smallest box containing $\mathcal{L}$ and let (c_i, d_i) be the extremal points of $\mathcal{L}$ chosen as above. Also choose $\alpha_0, \ldots, \alpha_{2m}$ as above with $\alpha_{2i+1} = a_0$. From proposition 3.5.2) $\mathcal{L}'$ is either contained inside this box or completely outside the box. We assume $\mathcal{L}'$ is inside the box. The line $A \times b_0$ has only two points on the $K^{-1}(0)$

contour these points must be (c_1, b_0) and (c_2, b_0). Hence $\mathcal{L}'$ is contained in one of $[a_1, a_0] \times B$ or $[a_0, a_2] \times B$. In fact from proposition 3.4.1) we can see that $\mathcal{L}$ must be contained in one of the *boxes* $[a_1, \alpha_{2i}] \times B$ or $[\alpha_{2i+2}, a_2] \times B$ and hence must have a different *associated point.*

We can not translate these results into a condition for a closed loop to occur. Choose two intervals $A = [a_1, a_2]$, $B = [b_1, b_2]$ such that $A \times B$ is a *box* in S and let (a_0, b_0) be its *associated point.* Now condition 1) of proposition 3.5.1 is equivalent to saying

1') $K(a_0, b_0) > 0$.

and condition 2) is equivalent to saying

2a') $K(a_1, b_0) < 0$,
2b') $K(a_2, b_0) < 0$,
2c') $K(a_0, b_1) < 0$,
2d') $K(a_0, b_2) < 0$.

If we check through all possible intervals to see if 1') and 2a')-2d') are all true then by proposition 3.5.1 we will have found all the boxes which contain such loops. This procedure may find several boxes which all contain the same loop. But we can refine it by using proposition 3.5.3).

The algorithm for finding all the boxes satisfying the above condition works as follows.

A) Loop through all points $(a_0, b_0) \in M \times N$ for which $K(a_0, b_0) > 0$. This satisfies condition 1') above.

A1) Label the points in $M \cup N$ $\alpha_0, \ldots, \alpha_{2n+1}$ chosen so that $\alpha_{2n+1} = a_0$, $\alpha_{2i} \in M$, $\alpha_{2i+1} \in N$ and $\alpha_i < \alpha_{i+1}$.

B) First we check condition 2a' above. Let $k = 2n$.

B1) Is $K(\alpha_k, b_0) < 0$? If so condition 2a') is true and we skip forward to C).

B2) Choose a larger box, let $k = k - 2$ if $k = 0$ then we have tested all possible boxes and 2a') has failed we now return to A) above to try a different (a_0, b_0).

B3) Is $\kappa(\alpha_{k+1}) > \kappa(a_0)$? If so then f does not assume its largest value on the interval $[\alpha_k,\ ?\ ]$ at a_0 so the test has failed and we return to A).

B4) Loop back to B1).

C) Now check condition 2b') above, let $k = 0$.

C1) If $K(\alpha_k, b_0) < 0$ go to D).

C2) Let $k = k + 2$ if $k = 2n$ return to A).

C3) If $\kappa(\alpha_{k-1}) > \kappa(a_0)$ return to A).

C4) Loop back to C1).

Now label the points in $M \cup N$ $\beta_0, \ldots, \beta_{2n+1}$ chosen so that $\beta_{2n+1} = b_0$, $\beta_{2i} \in N$, $\beta_{2i+1} \in M$ and $\beta_i < \beta_{i+1}$.

D) Now check condition $2c'$) , let $k = 2n$.

D1) If $K(a_0, \beta_k) < 0$ go to E).

D2) Let $k = k - 2$ if $k = 0$ return to A).

D3) If $\kappa(\beta_{k-1}) < \kappa(b_0)$ return to A).

D4) Loop back to D1).

E) Now check condition $2d'$) , let $k = 0$.

E1) If $K(a_0, \beta_k) < 0$ go to F).

E2) Let $k = k + 2$ if $k = 2n$ return to A).

E3) If $\kappa(\beta_{k-1}) < \kappa(b_0)$ return to A).

E4) Loop back to E1).

F) All the conditions have now been satisfied.

A suitable starting point for drawing these missing pairs can be found by moving one point on the curve $y = \kappa(t)$ "left and upwards" from b_0 until a point c with $\kappa(c) = \kappa(a_0)$ is found. The routine for finding these points is run as normal starting in "Descend" at the pair (a_0, c). In this case the routine will stop when (a_0, c) is reached again. This will always happen in the routine "Ascend Left". Note that because the starting point and ending points are the same these pairs of points will give a closed loop in the Rotational Symmetry Set.

It may occur that one minimum and one maximum both have the same y-values. When this happens there will be an isolated pair which will give a single point on the RSS. A separate condition in the routine would find such pairs.

§3.6 Have any pairs been found twice?

We observe from proposition 3.4.1 that for a generic curve the $K^{-1}(0)$ contour has no self intersections away from the line $\{t_1 = t_2\}$. So from any point on the contour either a) the point is on a closed loop or b) the point is on part of the contour which starts and ends at extrema (on the line $\{t_1 = t_2\}$). In case a) we use proposition 3.5.3 to show that the point is not duplicated. For case b) we examine the sign of K on either side of the contour and from proposition 3.4.1 15) and 16)

we see that one of end point must be a maximum of κ and the other must be a minimum. Hence such points will always be reached once and only once by the algorithm.

§3.7 Finding the RSS part 2

The problem of finding part 2 of the Rotational Symmetry Set is very similar to finding part 1. We wish to find the set

$$\{(L,R) \in I \times I \mid \kappa(L) = -\kappa(R)\}.$$

By locally reversing orientation as in §2.3 this is the same as finding the set

$$\{(L,R) \in I \times I \mid \kappa(L) = \overline{\kappa}(R)\}$$

where $\overline{\kappa}(R) = -\kappa(R)$. Apart from the ending points this is locally the same as the problem in the previous section of this chapter. As before four routines are used to control the movement (fig 3.14). These are

"Shear Left"	Left point moves up and left, right point moves down and right
"Shear Right"	Left point moves up and right, right point moves down and left
"Hump Left"	Left point moves down and left, right point moves up and left.
"Hump Right"	Left point moves down and right,right point moves up and right.

Note the sign of $\kappa(L)$ or $\kappa(R)$ does not affect the type of movement. The movement routines start and stop at the zeros of κ and it is also possible to have closed loops appearing. The details of finding the second part of the RSS have been omitted here but have been implemented in the program.

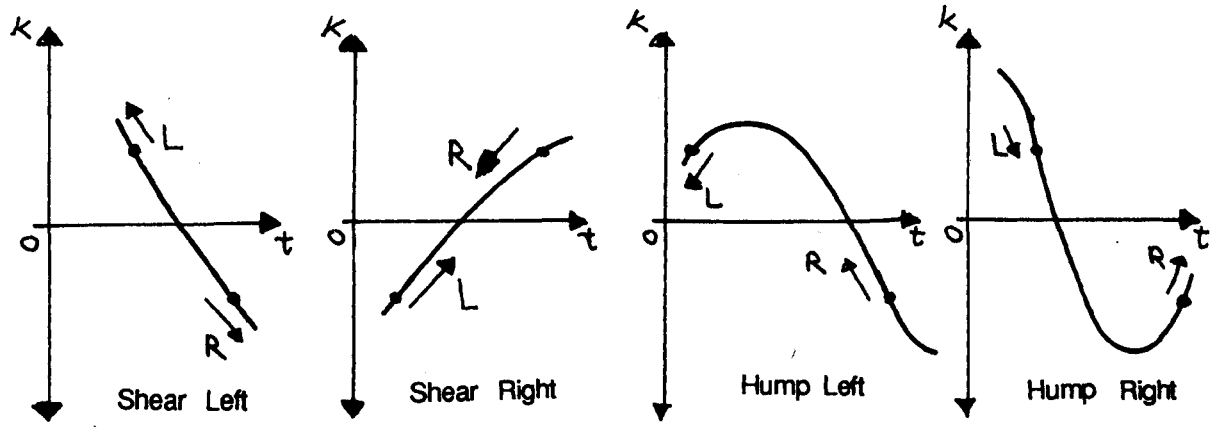


fig. 3.14) The different types of movment for calculating part 2 of the Rotational Symmetry Set.

Chapter 4: Calculating the Symmetry Set

One of the main reasons for developing a routine to find the RSS was because the routine could be modified to find the Symmetry Set. Whereas for the RSS we wanted to find the set

$$\{ (s,t) \in I \times I \mid \kappa(s) = \kappa(t) \}$$
$$= \{ (s,t) \in I \times I \mid K(s,t) = 0 \}$$

where K is the function $\kappa(s) - \kappa(t)$, for the Symmetry Set we need to find the set

$$\{ (s,t) \in I \times I \mid h(s,t) = 0 \}, \tag{4.0.1}$$

where h is the function defined below. This function cannot be split into the sum of two functions of one variable so the algorithm will be a generalization of that used in the previous chapter. Once the set 4.0.1) is found both the Symmetry Set and Midpoint Locus can be easily found.

In this chapter we describe the algorithm, which was developed during the project, to draw the SS. In §4.1 we describe a function $h : I \times I \to \mathbf{R}$ which picks out the SS. The main algorithm, for following the zero contour of this function is described in §4.2. Section 4.3 illustrates how the routine works in a few cases. The problems associated with the orientation of the bi-tangent circle are dealt with in §4.4. In §4.5 we describe how a set of starting points can be found so that the entire SS is found.

Let $\gamma : I \to \mathbf{R}^2$ be a parametrization of a closed smooth plane curve with tangent vector $T(t)$ and normal $N(t)$ at each point $\gamma(t)$ of the curve.

§4.1 A function h which defines the Symmetry Set

The first step in finding the SS is to find a function $h : I \times I \to \mathbf{R}$ where $h^{-1}(0)$ corresponds to the SS^{-1}. Several such functions have been described previously. In [Giblin-Brassett] the function

$$B(s,t) = (\gamma(s) - \gamma(t)) \cdot (T(s) - T(t)),$$

is used to study the local structure of the SS. A normalized version, $A(s,t)$, of this function is used in in [Dobkin-Thurston]. The normalization solves some computational problems. However a number of problems are encountered with both function.

1) If $T(s_1) = T(t_1)$ then $B(s_1,t_1) = 0$, but (s_1,t_1) does not necessarily lie on the SS^{-1}. Furthermore if we look at two nearby points (s_0,t_0), (s_2,t_2) then

$(T(s_0) - T(t_0))$ and $(T(s_2) - T(t_2))$ may point in almost opposite directions, causing a sign change in both functions. As the routine looks for changes in sign this will cause problems in the implementation of the routine.

2) We wish to study both the SS_1^{-1} and the SS_2^{-1}. The above functions only find SS_1^{-1}.

3) Numerical problems are encountered in both functions when $s \simeq t$. The function B has a high order of zero here, the normalization of A reduces this but numerical problems still remain.

The function h described below solves all the above problems.

The centre maps $C_1, C_2 : I \times I \to \mathbf{R}^2$ give the unique points $C_1(s,t)$, $C_2(s,t)$ about which there is a rotation which sends $\gamma(s)$ to $\gamma(t)$ and $T(s)$ to $\pm T(t)$. Let θ_1 be the angle of the rotation about $C_1(s,t)$ and θ_2 be the angle of the rotation about $C_2(s,t)$. Also let ϕ_1 be the angle between $V_1 = C_1(s,t) - \gamma(s)$ and $T(s)$, measured anticlockwise, which is the angle between $C_1(s,t) - \gamma(t)$ and $T(t)$. Let ϕ_2 be the angle between and $V_2 = C_2(s,t) - \gamma(s)$ and $T(s)$, which is also the angle between

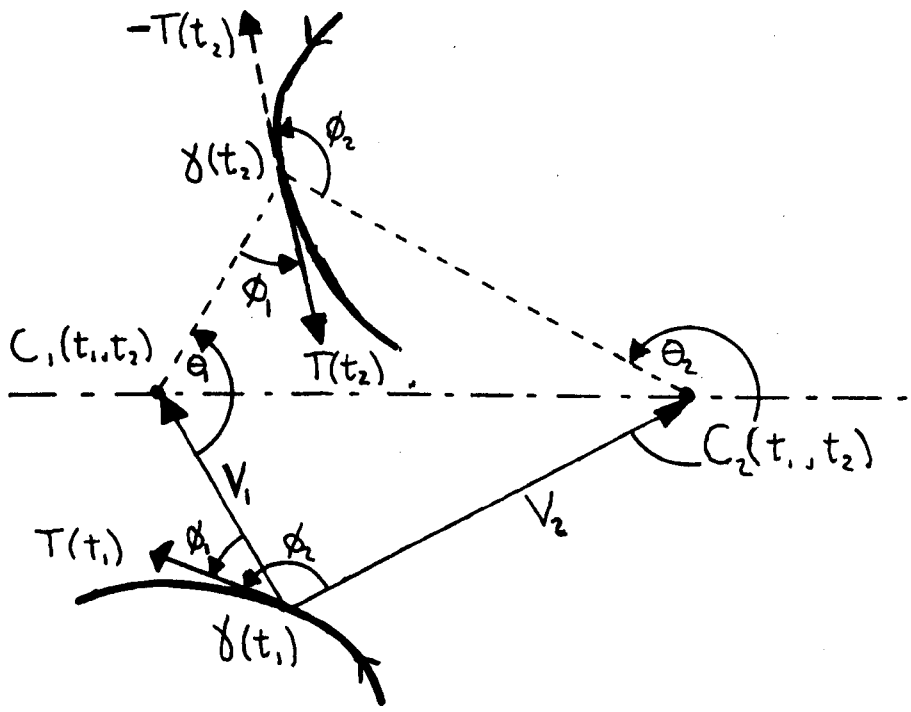


fig. 4.1) The geometry of the centre maps.

$C_2(s,t) - \gamma(t)$ and $-T(t)$. See fig. 4.1. Now by definition of the centre map (Def.

2.2.1.)

$$T(t) = T(s)\ e^{i\theta_1}$$
$$-T(t) = T(s)\ e^{i\theta_2}$$

so

$$e^{i\theta_1} = -e^{i\theta_2}$$
$$\Rightarrow \qquad \theta_1 - \theta_2 = \pi(\text{mod } 2\pi).$$

From the symmetry of the diagram we find that the angle between V_1, V_2 is $\pm\pi/2$, hence

$$\phi_2 = \phi_1 \pm \pi/2(\text{mod } 2\pi).$$

There will be bi-tangent circles whenever the tangent vector $T(s)$ is at right angles to one of the two vectors V_1, V_2. It now follows that if $\phi_1 = \pi/2$ or $3\pi/2$ then $(s,t) \in SS_1^{-1}$, and if $\phi_1 = 0$ or π then $\phi_2 = \pi/2$ or $3\pi/2$ so $(s,t) \in SS_2^{-1}$.

We now define a function $h_1 : I \times I \to \mathbf{R}$ by

$$h_1(s,t) = \sin(2\phi_1)$$

which is zero whenever there is a bi-tangent circle. This function is well defined provided $T(s) \neq T(t)$.

Both points $C_1(s,t)$ and $C_2(s,t)$ lie on the perpendicular bisector of the line joining $\gamma(s)$ and $\gamma(t)$. As the angle between $T(s)$,$T(t)$ tends to zero $C_1(s,t)$ will move along this line towards infinity and $C_2(s,t)$ will move towards $\frac{1}{2}(\gamma(s) + \gamma(t))$ (fig. 4.2). The function $h_2 : I \times I \to \mathbf{R}$ defined by

$$h_2(s,t) = -\sin(2\phi_2)$$

has the same value as h_1 whenever $T(s) \neq \pm T(t)$. But it is defined when $T(s) = T(t)$ and $\gamma(s) \neq \gamma(t)$. So using these two functions we can define $h_3 : I \times I \to \mathbf{R}$ by

$$h_3(s,t) = \begin{cases} h_1(s,t) & T(s) \neq T(t), \\ h_2(s,t) & T(s) = T(t),\ s \neq t. \end{cases}$$

We are only interested in the angle ϕ_2 modulo π so no problem is encountered as C_1 passes through infinity and angle is changed by π radians.

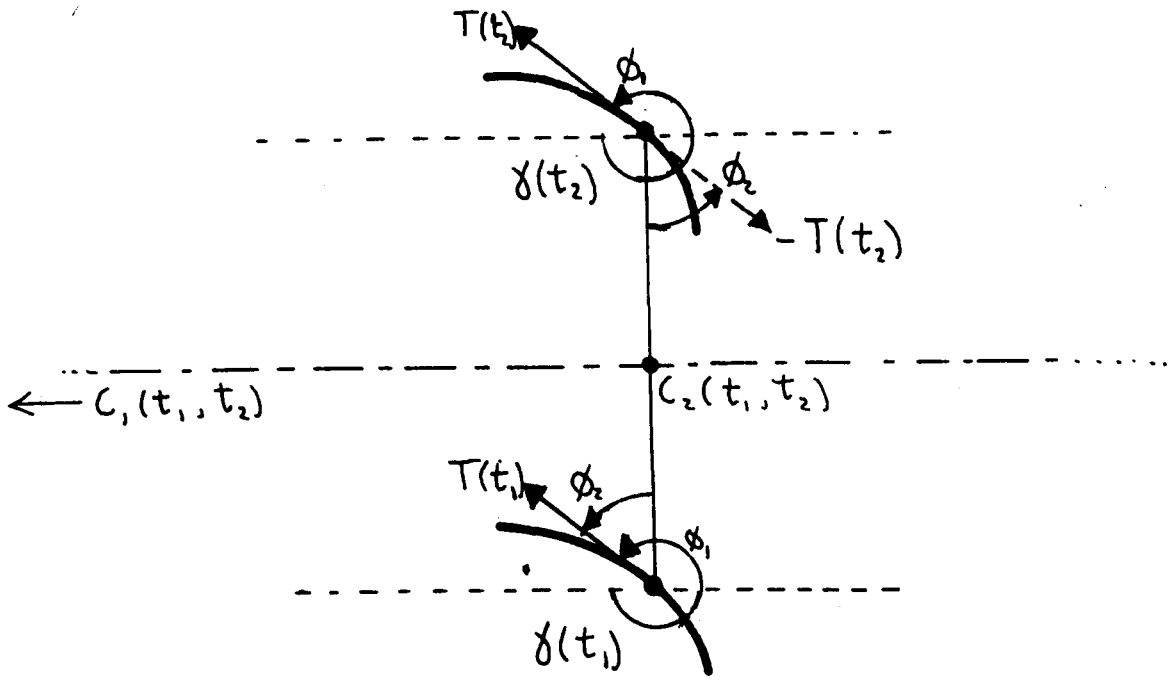


fig. 4.2) The situation when the tangents are parallel.

Neither h_1 or h_2 is defined for $s = t$ and near $s = t$ there are numerical problems in calculating these functions. When $s = t$ then the condition for (s,t) to be in SS_1^{-1} is that $\kappa'(s) = 0$ so this is the function we use on this region. For $|s-t|$ small we interpolate between the function h_1 and κ'. Let $\lambda = |s-t|$ and choose some small tolerance $\eta > 0$. For $\lambda < \eta$ we use the function $h_4 : I \times I \to \mathbf{R}$ defined by

$$h_4(s,t) = \lambda h_1(s,t) + (\eta - \lambda)\kappa'(\frac{1}{2}(s+t)),$$

where κ is the curvature. We now can define a function $h : I \times I \to \mathbf{R}$ for all (s,t).

Lemma 4.1.1

Let

$$h(s,t) = \begin{cases} h_1(s,t) = \sin(2\phi_1) & \text{if } T(s) \neq T(t),\ |s-t| > \eta\ , \\ h_2(s,t) = -\sin(2\phi_2) & \text{if } T(s) = T(t),\ |s-t| > \eta, \\ h_4(s,t) & \text{if } 0 < |s-t| \leq \eta \\ \quad = \lambda h_1 + (\eta - \lambda)\kappa'(\frac{1}{2}(s+t)) & \\ \kappa'(s) & \text{if } s = t. \end{cases}$$

If $|s-t| > \eta$ or $s = t$ then $h^{-1}(0)$ is the set $SS_1^{-1} \cap SS_2^{-1}$. If $0 < |s-t| < \eta$ then $h^{-1}(0)$ is an approximation to $SS_1^{-1} \cap SS_2^{-1}$. two sets are identical. We note that this function is symmetric.

§4.2 Following the $h^{-1}(0)$ contour.

The next stage in calculating the symmetry set is to find an algorithm for following the $h^{-1}(0)$ contour i.e. given a pair $(s,t) \in \mathbf{R}^2$ for which $h(s,t) = 0$ find another pair (s_1, t_1) nearby for which $h(s_1, t_1) = 0$. Additional information specific to the Symmetry Set will help improve the algorithm.

Let $s, t \in I$ be two points such that $h(s,t) = 0$ and the curve at points $\gamma(s)$ and at $\gamma(t)$ is oriented anticlockwise round the bitangent circle which we shall call $B(s,t)$. We also write $r_{B(s,t)}$ for its radius and $\kappa_{B(s,t)}$ for its curvature. These quantities are always positive. The bi-tangent circle is said to be **osculating** at $\gamma(s)$ if $\kappa(s) = \kappa_{B(s,t)}$ see §2.1.

The following results about the local structure of the h^{-1} contour come from [Giblin-Brassett] propositions 2.2 and 2.3 and are obtained by considering the partial derivatives of a function $g(s,t) = (\gamma(s) - \gamma(t)) \cdot (T(s) - T(t))$. When the curve is oriented anticlockwise round the circle at both points of contact the zero set of this function is the same as $h^{-1}(0)$.

Proposition 4.2.1 [Giblin-Brassett]

Figure 4.3 illustrates the local structure of the $h^{-1}(0)$ contour for a generic curve in the (s,t) plane (when γ is oriented anticlockwise round $B(s,t)$ at both points of contact) in the following cases.

1) $B(s,t)$ not osculating at $\gamma(s)$ or $\gamma(t)$ and $\kappa_B(s,t)$ between $\kappa(s)$ and $\kappa(t)$.
2) $B(s,t)$ not osculating at $\gamma(s)$ or $\gamma(t)$ and $\kappa_B(s,t)$ not between $\kappa(s)$ and $\kappa(t)$.
3) $B(s,t)$ osculating at $\gamma(s)$ but not at $\gamma(t)$ and $\kappa'(s)(1 - r_{B(s,t)}\ \kappa(t)) > 0$.
4) $B(s,t)$ osculating at $\gamma(s)$ but not at $\gamma(t)$ and $\kappa'(s)(1 - r_{B(s,t)}\ \kappa(t)) < 0$.
5) $B(s,t)$ osculating at $\gamma(t)$ but not at $\gamma(s)$ and $\kappa'(t)(1 - r_{B(s,t)}\ \kappa(s)) < 0$.
6) $B(s,t)$ osculating at $\gamma(t)$ but not at $\gamma(s)$ and $\kappa'(t)(1 - r_{B(s,t)}\ \kappa(s)) > 0$.
7) An isolated point when $B(s,t)$ osculates at both $\gamma(s)$ and $\gamma(t)$ and $\kappa'(s)\ \kappa'(t) > 0$.
8) Two smooth curves crossing when $B(s,t)$ osculates at $\gamma(s)$ and $\gamma(t)$ and $\kappa'(s)\ \kappa'(t) < 0$.
9) $s = t$ and $\kappa'(s) = 0$ i.e. at a vertex of the curve.

Cases 7) and 8) above will not occur for a generic curve but may occur in a generic family of curves. The associated transitions on $h^{-1}(0)$ are shown in figure 4.4.

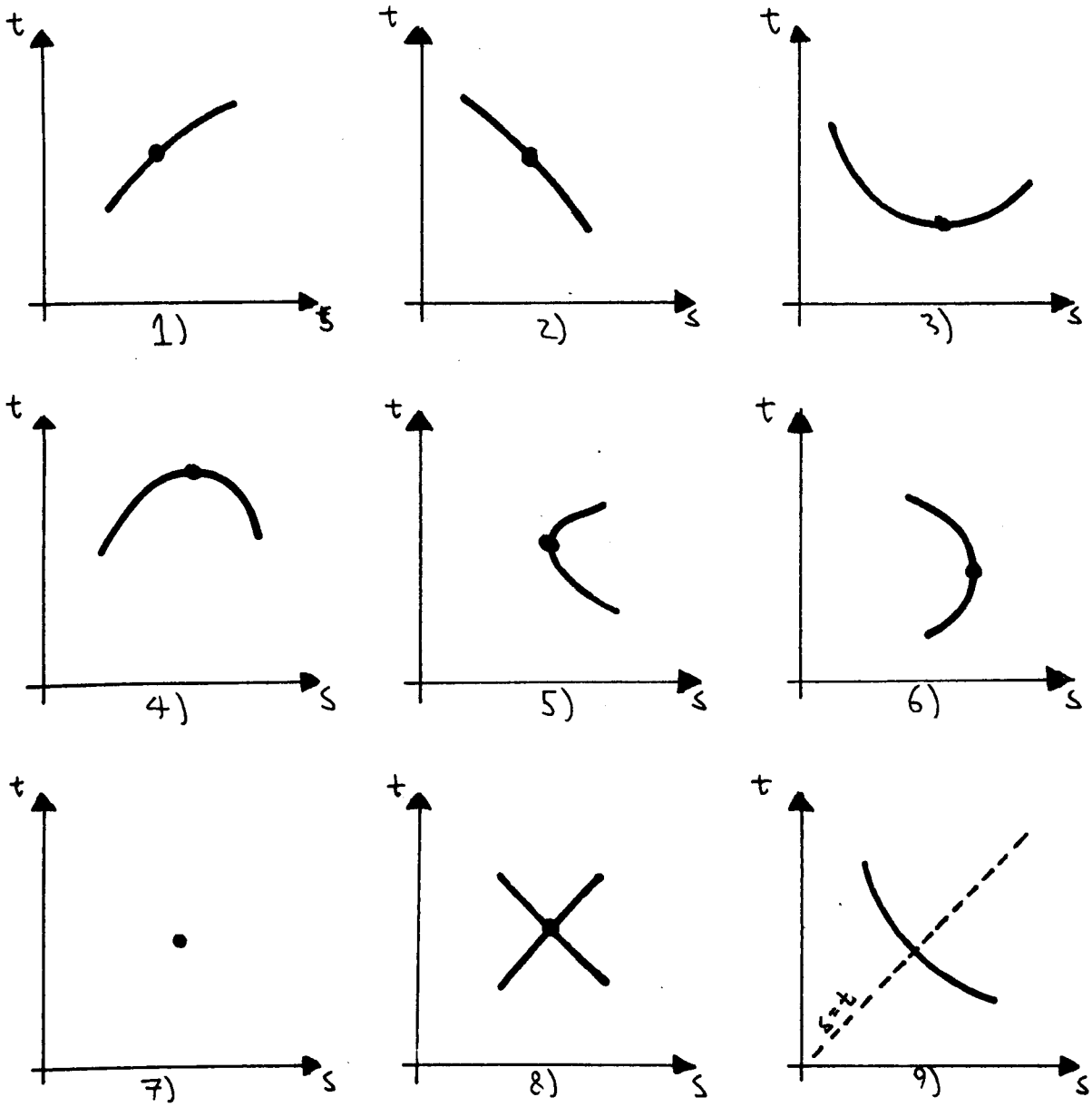


fig. 4.3) The local structure of the $h^{-1}(0)$ contour.

The algorithm must be able to cope with cases 1)–6) but need not be able to cope with the bi-osculating cases. However it has to work for curves which nearly satisfy the bi-osculating condition as it is pictures of these very transitions we are interested in.

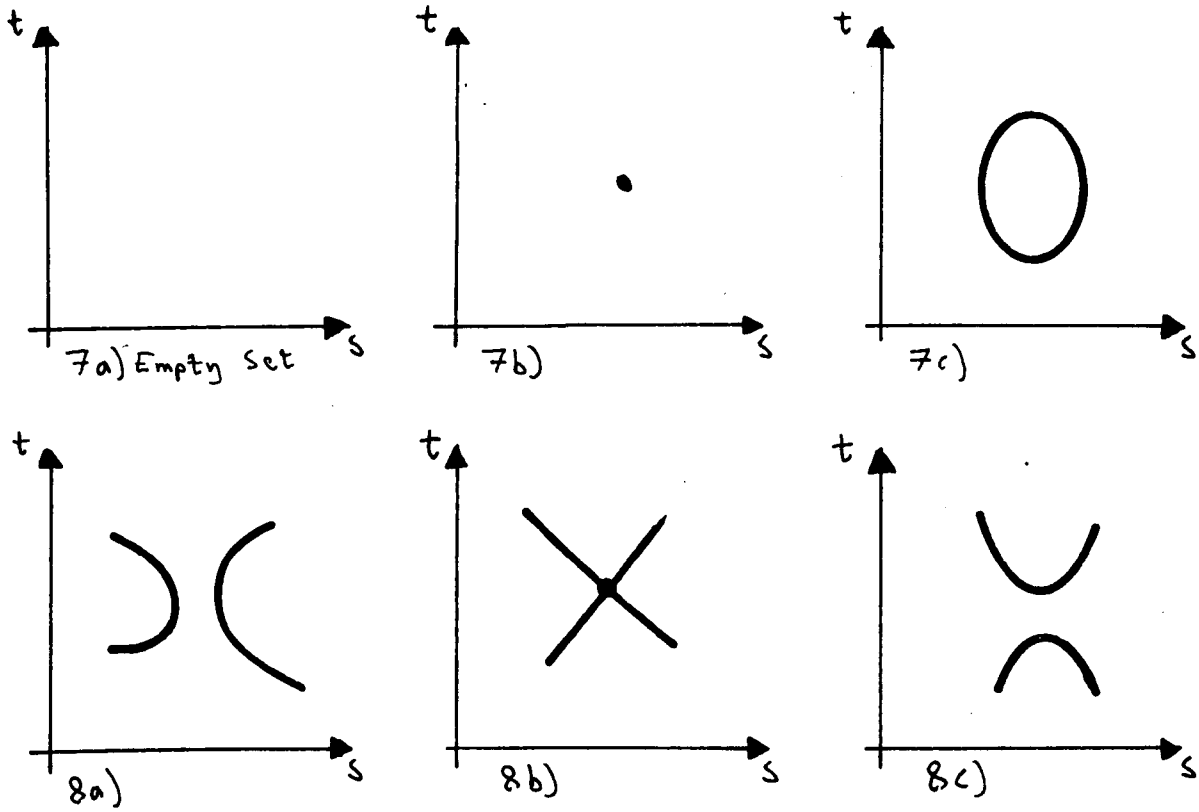


fig. 4.4) The transitions in the bi-osculating case.

Using the above information we can now devise an algorithm for following the $h^{-1}(0)$ contour. Assume we have a pair (s_0, t_0) such that $h(s_0, t_0) = 0$ and we know the direction of movement in the s and t directions. This information is stored in two variables $dirS$, $dirT$ with $dirS = +1$ if the value of s is increasing and $dirS = -1$ if s is decreasing. Let inc be the maximum allowable difference between two consecutive values of s or t. We now let $ds = dirS \times inc$ and $dt = dirS \times inc$.

4.2.2: Step A

To find the next pair of points we consider a box with corners at (s_0, t_0), $(s_0 + ds, t_0)$, $(s_0 + ds, t_0 + dt)$ and $(s_0, t_0 + dt)$. If we are not near an osculating point, i.e. in cases 1) or 2) of proposition 4.2.1, then the $h^{-1}(0)$ contour will pass through the point (s_0, t_0) and somewhere on the perimeter of the box between $(s_0 + ds, t_0)$ and $(s_0 + ds, t_0 + dt)$ or between $(s_0 + ds, t_0 + dt)$ and (s_0, t_0) (fig 4.5).

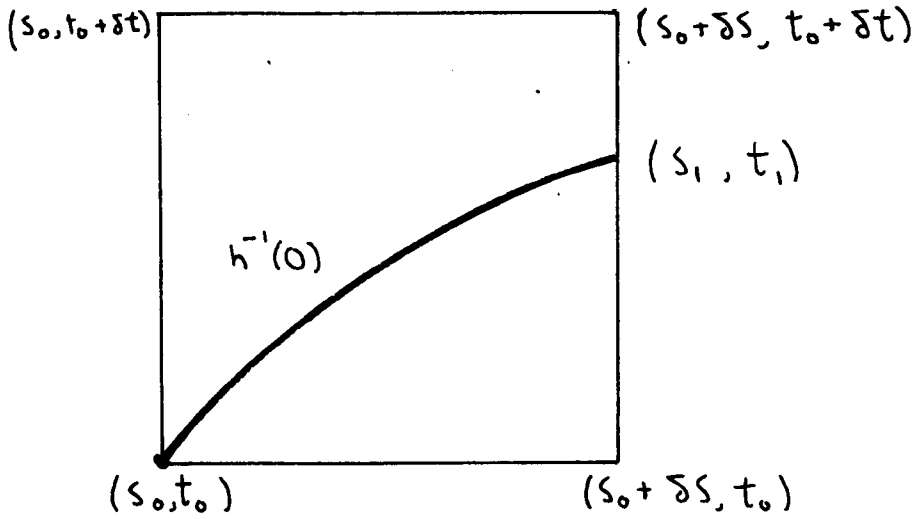


fig. 4.5) Finding the next point on the perimeter of the box.

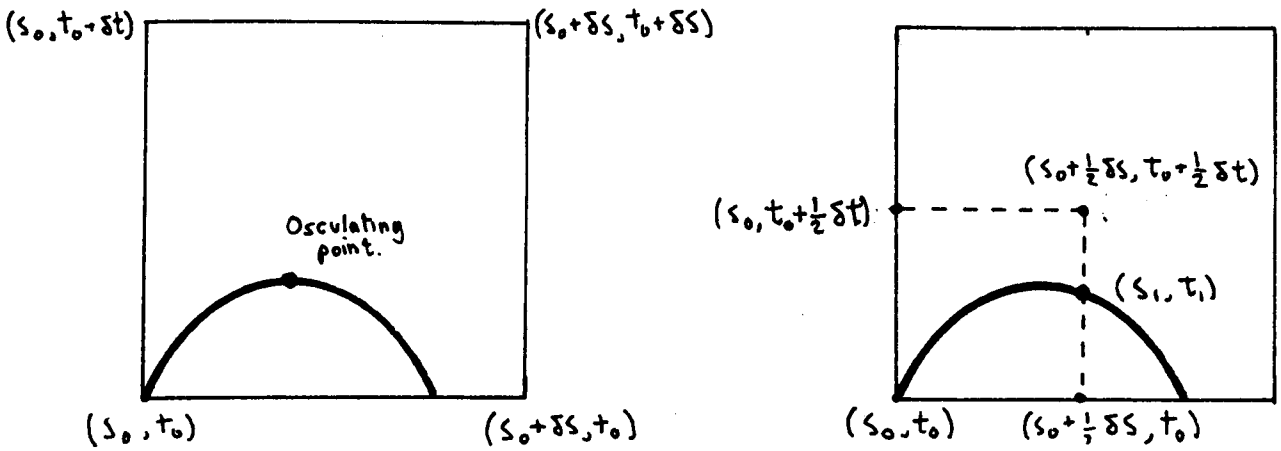


fig. 4.6) Reducing the size of the box to find a solution in the osculating case.

A simple routine can be used to converge to the solution of $h(s,t) = 0$ on one of these two sides. Let this solution be (s_1, t_1).

In the osculating case the $h^{-1}(0)$ contour does not necessarily cross either of these two sides. If this happens we reduce ds, dt by half and consider a smaller box (fig 4.6). This size reduction repeats until a solution is found.

In one of the cases 8a) or 8c) the solution found need not be on the same part of the $h^{-1}(0)$ contour. Two ways in which this can happen as shown in figure 4.7. Therefore we need to find some method of finding out whether the points are connected, also it would be useful to find out whether the contour changes direction

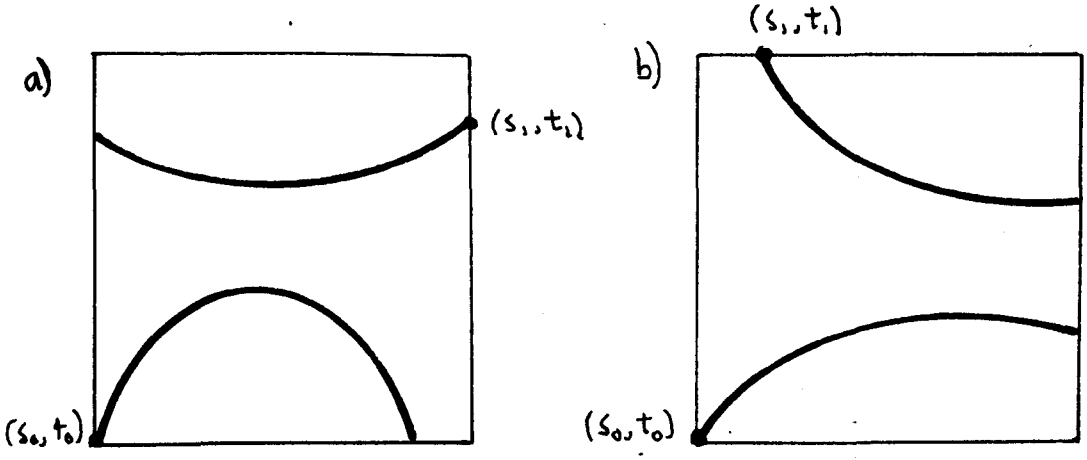


fig. 4.7) Two situations where the wrong point is found.

i.e. if osculation occurs (cases 3)–6) of proposition 4.2.1). If

$$(\kappa_B(s_0,t_0)-\kappa(s_0))(\kappa_B(s_1,t_1)-\kappa(s_1))<0 \tag{4.2.1}$$

then there is possibly a pair (s,t) on $h^{-1}(0)$ with $B(s,t)$ osculating at $\gamma(s)$ and if

$$(\kappa_B(s_0,t_0)-\kappa(t_0))(\kappa_B(s_1,t_1)-\kappa(t_1))<0 \tag{4.2.2}$$

then there may be a pair with $B(s,t)$ osculating at $\gamma(t)$.

To see which of these conditions hold in figure 4.7b) we study what happens in the A_2^2 transition which is shown in figure 4.8. As we change the curve through the bi-osculating case the signs of $\kappa_B(s_0,t_0)-\kappa(t_0)$ and $\kappa_B(s_1,t_1)-\kappa(t_1)$ will not change (provided the box is small enough). After the transition (fig 4.8b) equation 4.2.2 holds between (s_0',t_0') and (s_1',t_1') therefore equation 4.2.2 also holds before the transition between (s_0,t_0) and (s_1,t_1) (fig 4.8a). We can also show that equation 4.2.1 does not hold between the two points. By a similar argument we can see that both equations 4.2.1 and 4.2.2 hold between (s_0,t_0) and (s_1,t_1) in figure 4.7a) and in this case the two points are not connected either before or after the transition.

Once we have established if there is a possible osculating point we try to find the pair (s,t) for which osculation occurs. Assume that equation 4.2.1 holds between (s_0,t_0) and (s_1,t_1). First we try to find a pair (s,t) in the box with $s=\frac{1}{2}(s_0+s_1)$ and $h(s,t)=0$. If no such pair is found the program goes to (A) above. If the equation 4.2.1 holds between (s,t) and (s_1,t_1) we let $s_0:=s$, $t_0:=t$ and if it holds between (s_0,t_0) and (s,t) we let $s_1:=s$ and $t_1:=t$. If these tests both fail or both

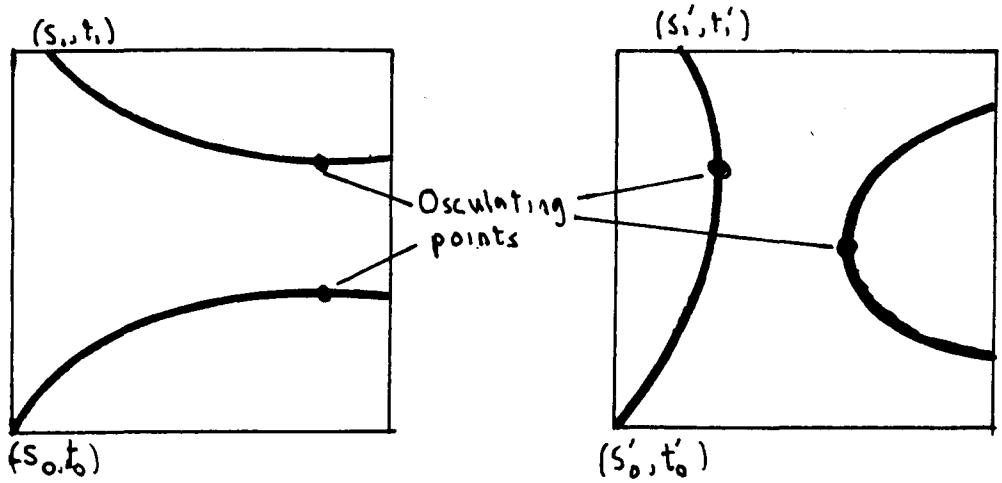


fig. 4.8) The effect of the bi-osculating transition on fig. 4.7b).

succeed then we try a smaller box by returning to (step A) above. This routine repeats for a fixed number of times until (s,t) is sufficiently close to an osculating pair. A final check is made using cases 3) and 4) of proposition 4.2.1 and considering the direction of movement, $dirT$, of the point. We check that

$$\kappa'(s)\ (1 - r_{B(s,t)}\kappa(t))\ dirT < 0; \qquad 4.2.3$$

if this fails the program will return to (A). Otherwise (s,t) become the next pair to start with and $dirT = -dirT$.

A similar process is used if equation 4.2.2 holds. If both equations hold we know that (s_0, t_0) and (s_1, t_1) are on different parts of the contour so we need to try a smaller box (return to (A)). If neither condition holds then there is no osculating pair and the direction of movement is unchanged.

The operation of the program is best illustrated by a flowchart fig. 4.9, fig. 4.10.

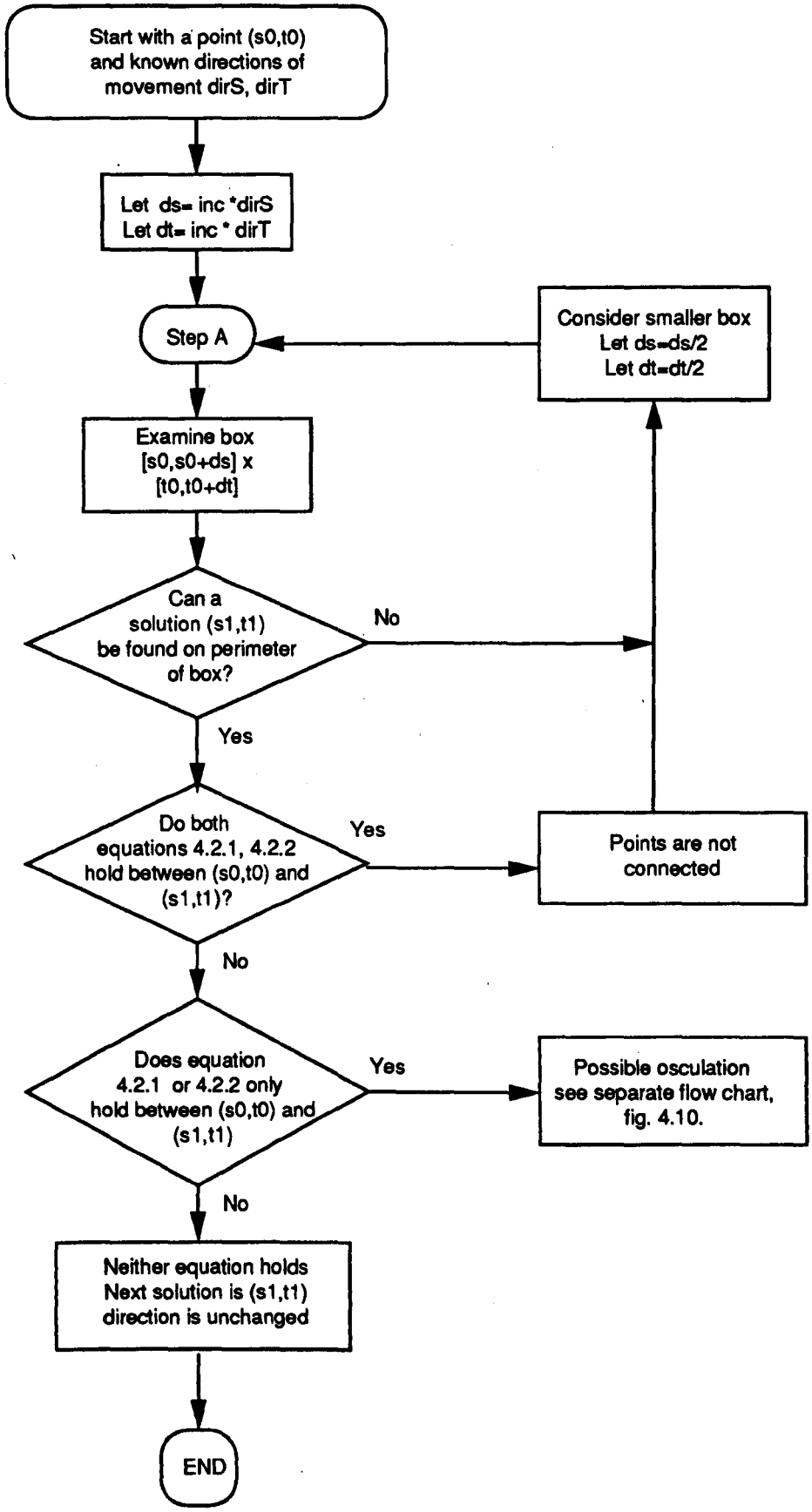


Figure 4.9 Flowchart for finding next pair of points

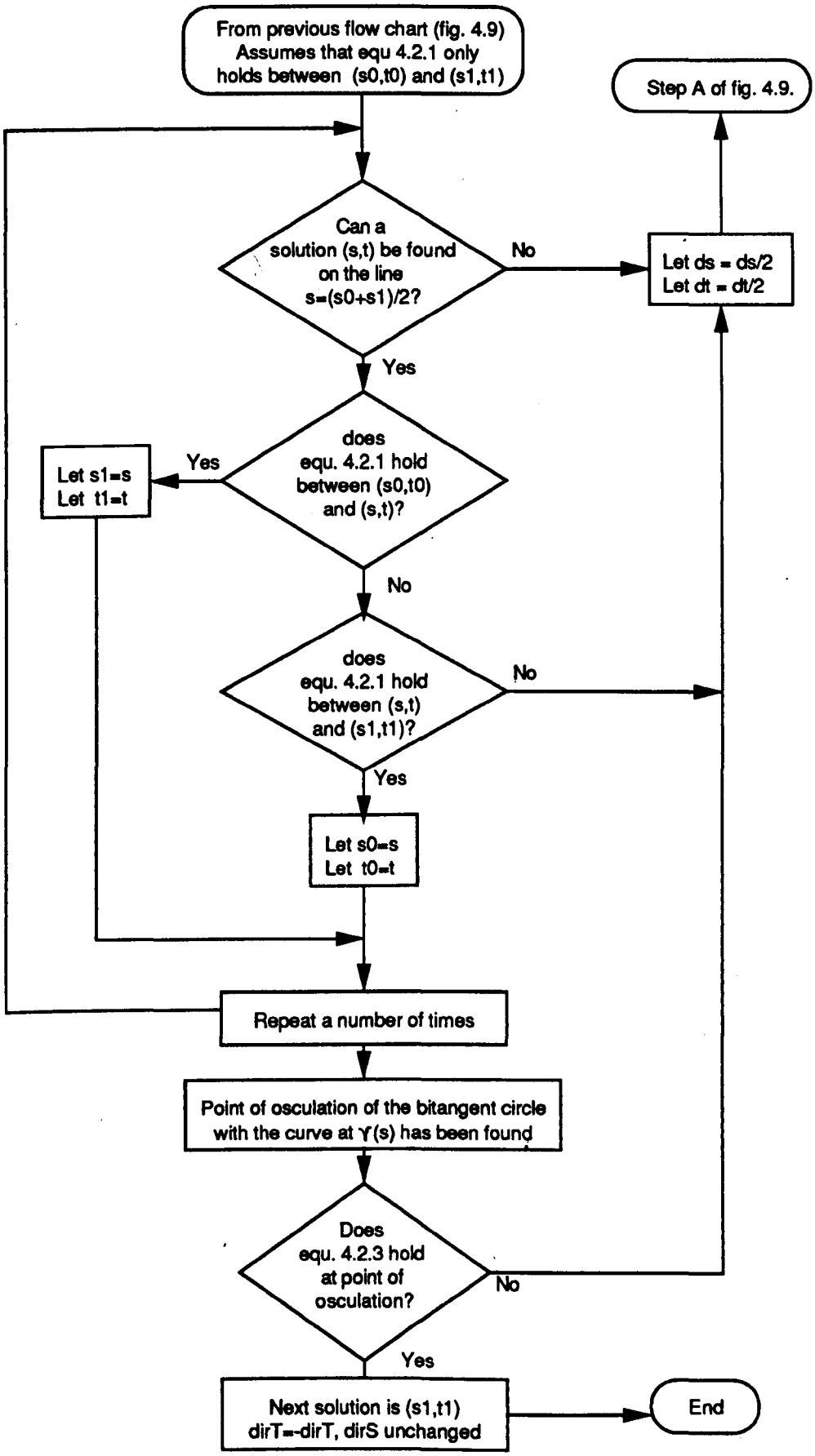


Fig 4.10 Flowchart for finding an osculating point when there is a possible osculation with $\gamma(s)$.

§4.3 Examples of the operation of the routine.

Example 1)

1a) Start at (s_0, t_0) with $dirS = +1$ and $dirT = -1$

1b) Find solution (s_1, t_1) on perimeter of box (fig. 4.11).

1c) Neither equation 4.2.1 or equation 4.2.2 holds so (s_1, t_1) is the next pair to start with.

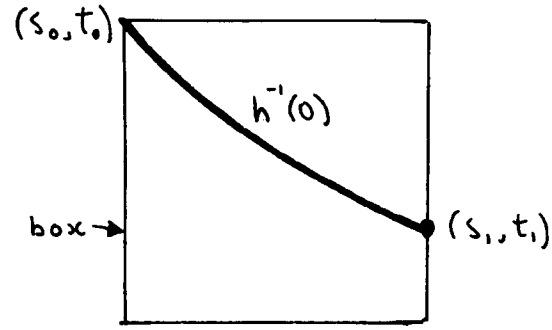


fig. 4.11) The operation of the routine in example 1).

Example 2)

2a) Start at (s_0, t_0) with $dirS = +1$, $dirT = -1$.

2b) Find solution (s_1, t_1) on perimeter of box (fig 4.12a).

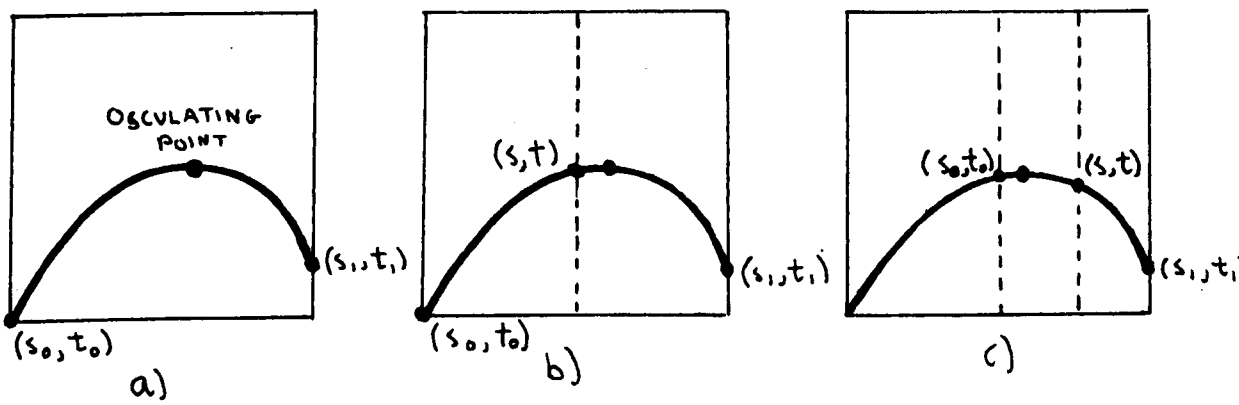


fig. 4.12) The operation of the routine in example 2).

2c) Equation 4.2.1 holds so we look for a possible osculating point:-
find solution (s, t) on line $s = \frac{1}{2}(s_0 + s_1)$ (fig. 4.12b).

2d) Equation 4.2.1 holds between (s, t) and (s_1, t_1):-
Let $(s_0, t_0) = (s, t)$;
Find solution (s, t) on line $s = \frac{1}{2}(s_0 + s_1)$ (fig. 4.12c).

2e) Repeat until we have converged to point of osculation.

2f) Test if $\kappa'(s)(1 - r_{B(s,t)}\kappa(t))\ dirT < 0$

Test passes so (s,t) is the next starting pair and $dirT = -dirT$.

Example 3

3a) Start at (s_0, t_0) with $dirS = +1$ and $dirT = -1$.

3b) Find solution (s_1, t_1) on perimeter of box (fig 4.13a).

3c) Equation 4.2.1 holds:-

try to find a solution (s,t) on line $s = \frac{1}{2}(s_0 + s_1)$ (fig. 4.13b).

3d) No solution found, consider smaller box (fig 4.13c).

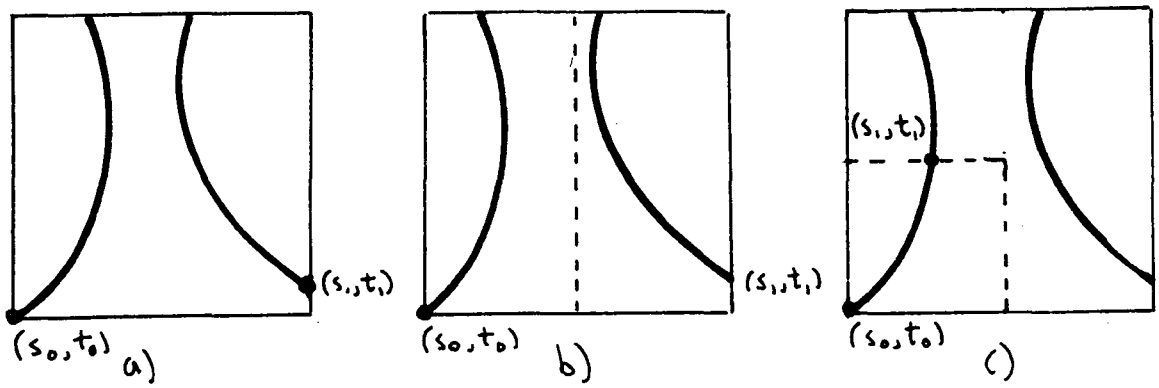


fig. 4.13) The operation of the routine in example 3).

§4.4 Dealing with orientation

Let γ_1 be a neighbourhood of $\gamma(s)$ in γ and γ_2 a neighbourhood of $\gamma(t)$. At present the routine only works when γ_1 , γ_2 are oriented anticlockwise round the circle . These orientations do not effect the $h^{-1}(0)$ contour but will change the results of proposition 4.2.1 and hence what happens in the osculating case. The four possibilities for the orientation are

1) γ_1, γ_2 oriented anticlockwise round the circle,
2) γ_1, γ_2 oriented clockwise round the circle,
3) γ_1 oriented clockwise and γ_2 oriented anticlockwise,
4) γ_1 oriented anticlockwise and γ_2 oriented clockwise

and are illustrated in figure 4.14.

The simplest way to overcome this problem is to let $\overline{\gamma}_1$, $\overline{\gamma}_2$ to be the curves γ_1 , γ_2 oriented anti-clockwise round the circle and then use the results from proposition 4.2.1. In practice once we have found a pair (s,t) for which $h(s,t) = 0$ we find the

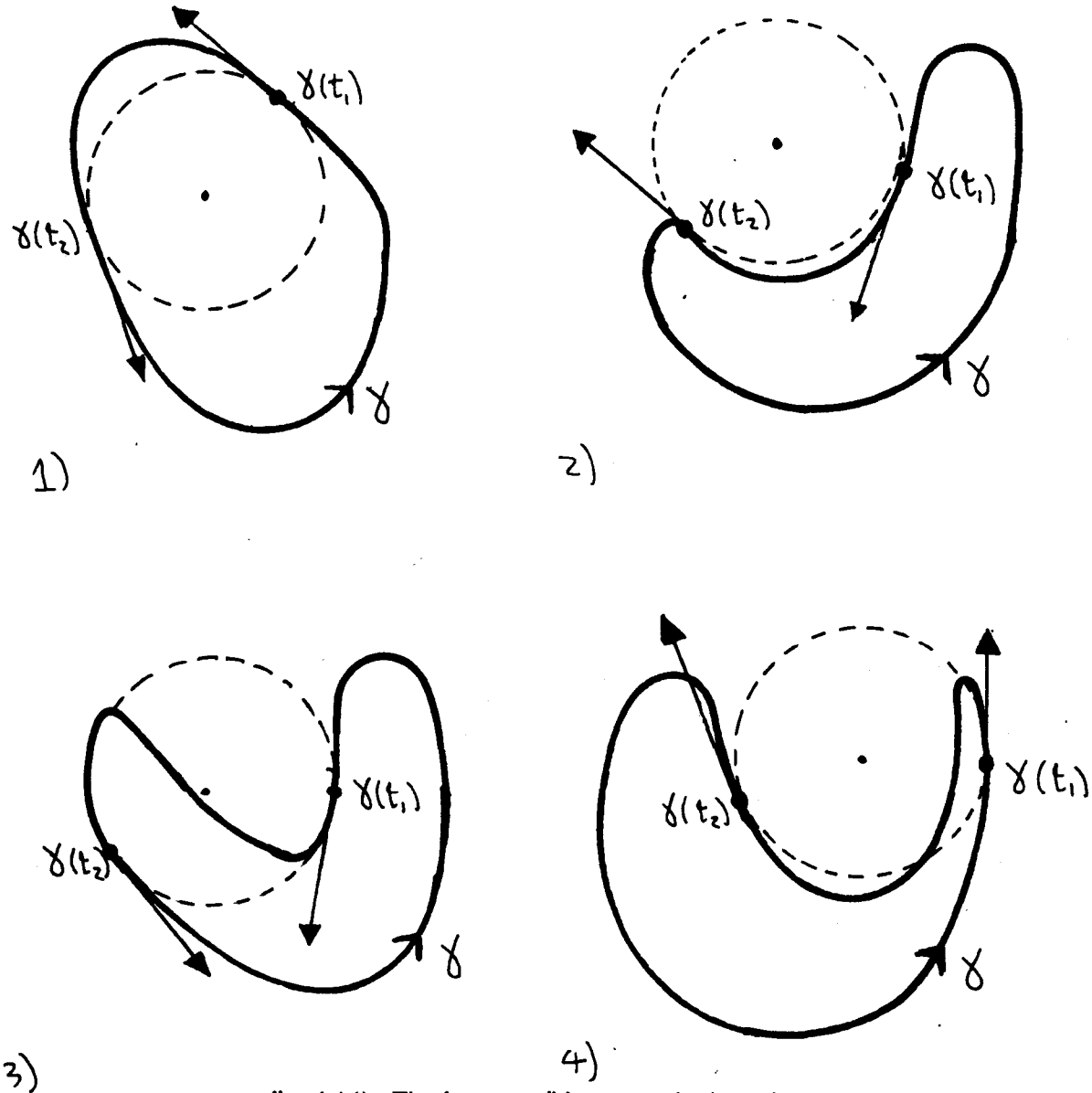


fig. 4.14) The four possible types of orientation.

centre of the circle and determine if $T(s)$ points anti-clockwise round the circle. If it does we let

$$\overline{T}(s) = T(s)$$
$$\overline{\kappa}(s) = \kappa(s)$$
$$\overline{\kappa'}(s) = \kappa'(s).$$

otherwise we let

$$\overline{T}(s) = -T(s)$$
$$\overline{\kappa}(s) = -\kappa(s)$$
$$\overline{\kappa'}(s) = +\kappa'(s).$$

We do the same for the orientation of $\overline{\gamma_2}$.

When following $h^{-1}(0)$ it is possible to change between orientations 1) and 2) or between 3) and 4). This happens when there is a bitangent line. In figure 4.15 we have orientation 1) applying for pair (s_1, t_1) , orientation 2) for pair (s_3, t_3) and

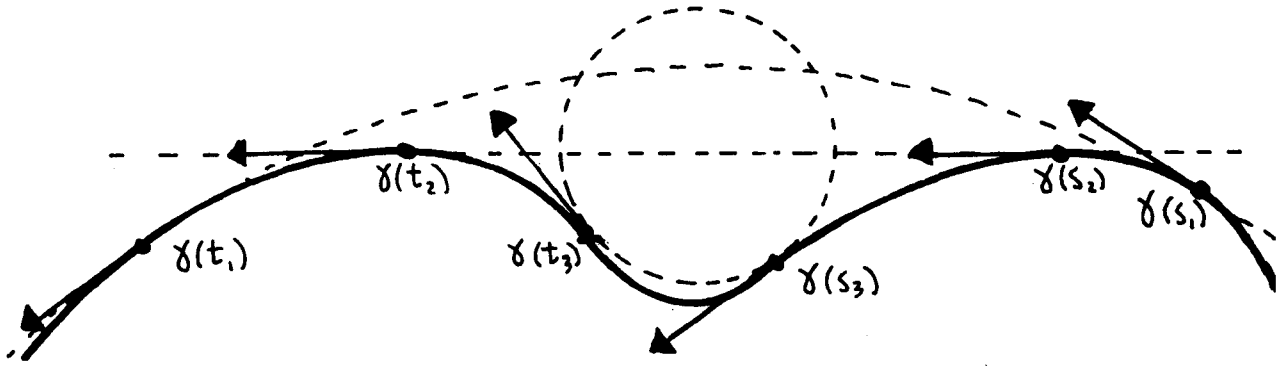


fig. 4.15) Change of orientation when there is a bi-tangent line.

for the pair (s_2, t_2) we have a bitangent line. By choosing the orientation at each point we can use the algorithm as above without having to worry about the global orientation. For a generic curve a bitangent line will not osculate at either point. Therefore when the tangents are parallel we do not test equations 4.2.1 or 4.2.2 so the routine will work for points like (s_2, t_2).

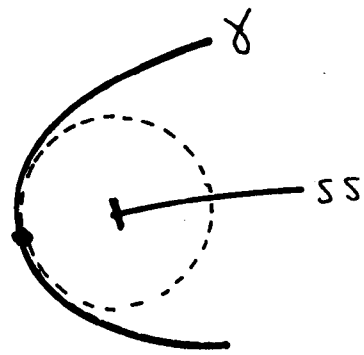
§4.5 Finding starting points

The final problem in calculating the symmetry set of a curve is to find suitable starting and ending points for following the contour. There are three different situations illustrated in figure 4.16:-

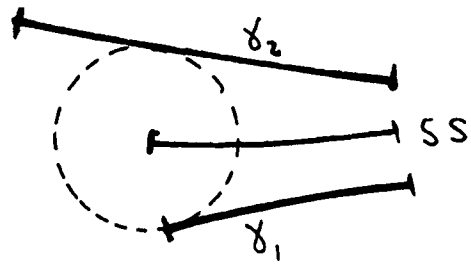
a) When $s = t$ and $\gamma(s)$ is a vertex of the curve;
b) When the curve consists of a number of disjoint segments:
   there may be a circle tangent at an end point of a segment
   and at some other point of any of the segments;
c) When the $h^{-1}(0)$ contour forms a closed loop e.g. in case 7c) of fig. 4.4.

To find all the starting points of type a) we just need to find the maxima and minima of curvature of the curve and store these in a list. We can start following the contour from each vertex in the list. When we reach the end of one segment of the contour the vertex we have arrived at is removed from the list to avoid redrawing this piece later on.

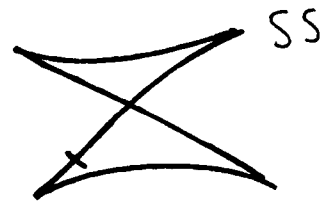
All the disjoint segments of a curve can be described by a single discontinuous



a) Vertex of curve.



b) End point of a disjoint segment of the curve.



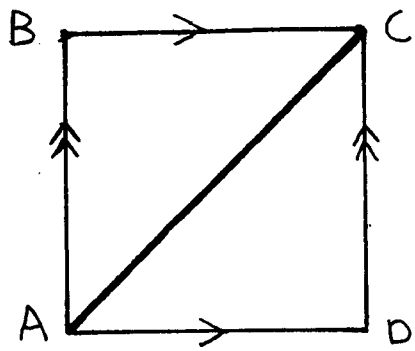
c) A closed loop on the Symmetry Set.

fig. 4.16) Starting points for following the Symmetry Set.

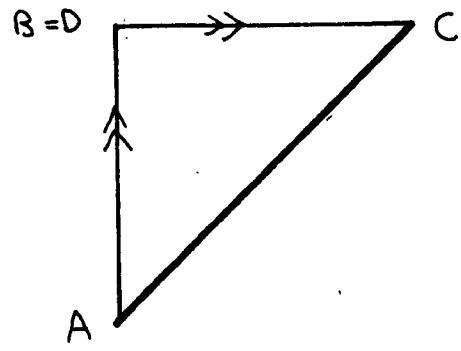
parametrization with one sub-interval of $[0,1]$ for each segment. For each pair of sub-intervals, $[s_0, s_1]$ and $[t_0, t_1]$ say, we examine the boundary of the box $[s_0, s_1] \times [t_0, t_1]$ to find the points where $h(s,t)$ vanishes. As for the vertices we make a list of all such points. We also need to consider when the two segments are the same and here we must include the vertices in the list as it is possible to start at the vertex and end along the boundary of the box. Checking through all pairs of segments will give a complete list. At each point of the list we must also find the values of $dirS$ and $dirT$ for starting the "following" routine.

For a closed curve the parameter can be thought of as lying in the unit sphere

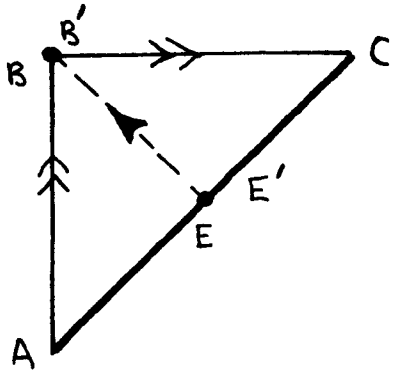
S^1. Topologically the set of unordered pairs (t_1, t_2) with $t_1, t_2 \in S^1$ is a Mobius



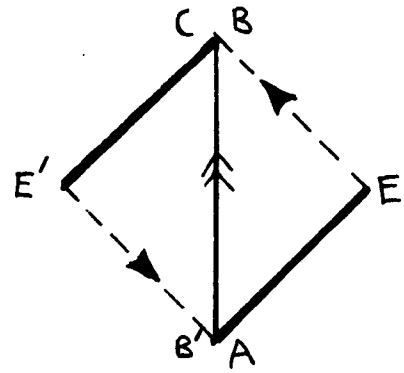
a) Opposite edges identified.



b) Lower and upper halves of triangle identified.



c) Cut along edge BE.



d) Joining identified edges to give a Mobius band.

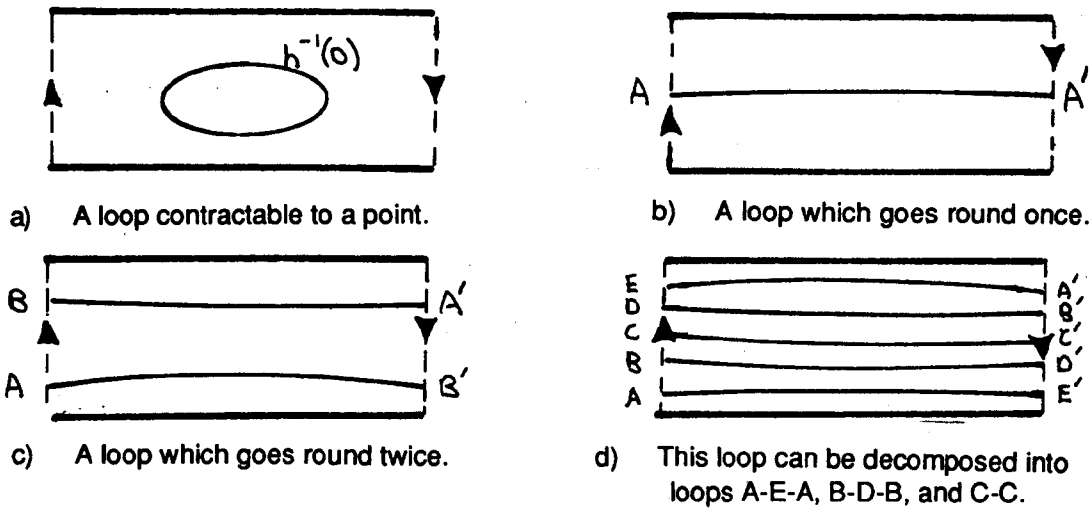
fig. 4.17) Showing how the set of unordered pairs (s,t) in $S^1 \times S^1$ is a Mobius band.

band (fig 4.17). There are three different kinds of simple closed loop which can occur on a Mobius band (fig 4.18):

1) Those that can be contracted to a point;
2) Those that go round the band once;
3) Those that go round the band twice.

Any loop which goes round the band more than twice is either the union of a number of loop of the types above or has a self-intersections.

We see that situation c) above splits into these three different cases. However the topology of a Mobius band prevents loops of type 2) being the zero contour of a generic function as the band can not be divided into positive and negative halves. For the other two cases we consider which points on the loop will give points which also lie on the Rotational Symmetry Set. A pair $(s,t) \in SS_1^{-1}$ for which $\kappa(s) = \kappa(t)$ will give rise to an inflection on both the Symmetry Set and the Rotational Symmetry Set (proposition 2.5.1). By calculating the RSS i.e. finding



a) A loop contractable to a point.

b) A loop which goes round once.

c) A loop which goes round twice.

d) This loop can be decomposed into loops A-E-A, B-D-B, and C-C.

fig. 4.18) The possible types of non-intersecting closed loop on a mobius band.

all pairs for which $\kappa(s) = \kappa(t)$ and then making a list of all the pairs for which $h(s,t) = 0$ we will know all the pairs on the $h^{-1}(0)$ contour for which $\kappa(s) = \kappa(t)$. From section 3.5 we know that the algorithm for the RSS will find *all* such possible pairs. We will now show that any closed loop on the SS contains a least one such pair. In case 1) we have the following lemma.

Lemma

If there exists a closed loop of type 1) above which is part of the set SS_1^{-1} then there exists some point (s,t) on the loop for which $\kappa(s)$ and $\kappa(t)$ are equal.

Proof

Parametrize the loop by $\mathcal{L} : I \to I \times I$ with $\mathcal{L}(\lambda) = (s(\lambda), t(\lambda))$. Let λ_0 , λ_2 be the values of the parameter with the lowest and highest values of $s(\lambda)$ and let λ_1 , λ_3 be the values which have the lowest and highest values of $t(\lambda)$ (fig 4.19). With out loss of generality we assume the loop is orientated anticlockwise and we have

$$0 = \lambda_0 < \lambda_1 < \lambda_2 < \lambda_3 < 1$$

as otherwise the loop would have self-intersections. Assume that $\gamma(s(\lambda_0))$ and $\gamma(t(\lambda_0))$ are oriented anticlockwise round the bitangent circle $B(\mathcal{L}(\lambda))$. Let

$$\kappa_1(\lambda) = \kappa(s(\lambda)) - \kappa_{B(\mathcal{L}(\lambda))}$$
$$\kappa_2(\lambda) = \kappa(t(\lambda)) - \kappa_{B(\mathcal{L}(\lambda))}$$

if γ is oriented anticlockwise round the circle $B(\mathcal{L}(\lambda))$ at $\gamma(s(\lambda))$ and $\gamma(t(\lambda))$, otherwise let

$$\kappa_1(\lambda) = \kappa(s(\lambda)) + \kappa_{B(\mathcal{L}(\lambda))}$$
$$\kappa_2(\lambda) = \kappa(t(\lambda)) + \kappa_{B(\mathcal{L}(\lambda))}.$$

These two functions are continuous as the change of orientation happens only when there is a bitangent line i.e. when $\kappa_{B(\mathcal{L}(\lambda))} = 0$. Each function is zero when the bitangent circle osculates at $\gamma(s(\lambda))$ or $\gamma(t(\lambda))$ respectively. We wish to show that there is some point for which these two functions are equal.

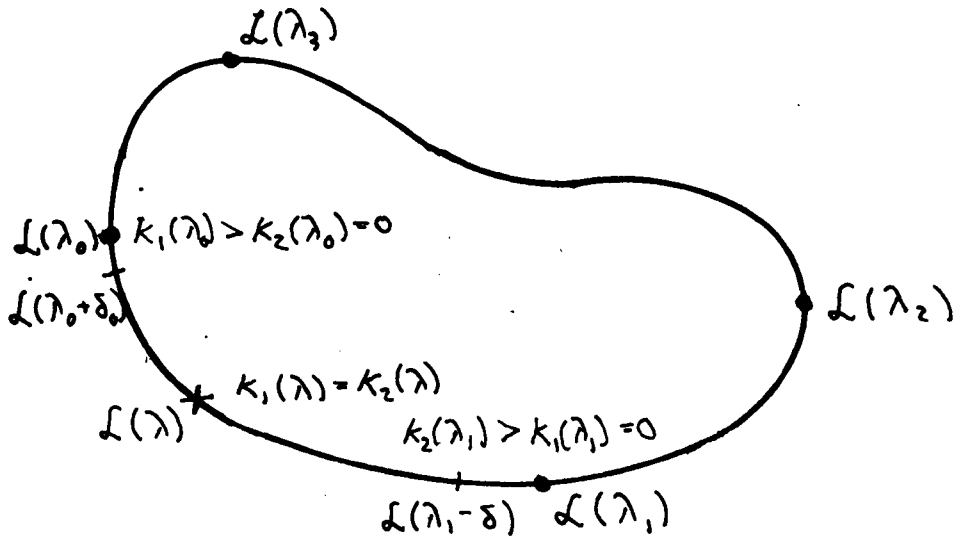


fig. 4.19) A point on a closed loop with equal values of curvature.

Let $s'(\lambda), t'(\lambda)$ denote the derivatives of s, t with respect to λ. From proposition 4.2.1 3-6) we have the following relations

$$s'(\lambda) = 0 \iff \kappa_2(\lambda) = 0$$
$$t'(\lambda) = 0 \iff \kappa_1(\lambda) = 0$$

In particular $\kappa_1(\lambda_1)$, $\kappa_2(\lambda_0)$ are both zero and $\kappa_1(\lambda_0)$, $\kappa_2(\lambda_1)$ are both non zero.

Assume $\kappa_1(\lambda_0)$ is positive. For a generic curve we can find a small $\delta_0 > 0$ such that neither $\kappa_1(\lambda)$ or $\kappa_2(\lambda)$ is zero for $\lambda \in (\lambda_0, \lambda_0 + \delta_0]$ also we can find a small $\delta_1 > 0$ such that neither $\kappa_1(\lambda)$ or $\kappa_2(\lambda)$ is zero for $\lambda \in [\lambda_1 - \delta_1, \lambda_1)$. By the choice of δ_0 we observe $\kappa_1(\lambda_0 + \delta_0) > 0$. Looking at the direction of the slope at $\mathcal{L}(\lambda_0 + \delta_0)$ we are in case 2) of proposition 4.2.1 so $\kappa_2(\lambda_0 + \delta_o)$ is also positive.

As $s'(\lambda_0 + \delta_0) > 0$ and $s'(\lambda_1 - \delta_1) > 0$ the number of zeros of $s'(\lambda)$ and hence $\kappa_2(\lambda)$ is even for $\lambda \in [\lambda_0 + \delta_0, \lambda_1 - \delta_1]$. Similarly the number of zeros of $\kappa_2(\lambda)$ is even. This means that $\kappa_2(\lambda_1 - \delta_1) > 0$ and from the choice of δ_1 we see that $\kappa_2(\lambda_1)$ is positive.

Finally we have $\kappa_1(\lambda_0) - \kappa_2(\lambda_0) > 0$ and $\kappa_1(\lambda_1) - \kappa_2(\lambda_1) < 0$ so there is some $\lambda \in [\lambda_0, \lambda_1]$ for which

$$\kappa_1(\lambda) = \kappa_2(\lambda)$$
$$\iff \kappa(s(\lambda)) = \kappa(t(\lambda)).$$

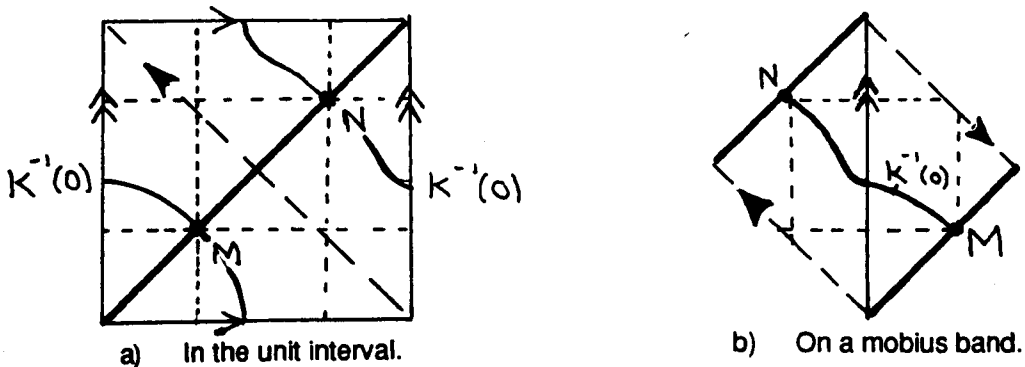
Corollary

Any segment of the $h^{-1}(0)$ contour which is a closed loop of type 1) above will have some point on it which gives an inflection on the Symmetry Set. This point can be found by calculating parts 1) and 2) of the Rotational Symmetry Set.

Proof

The assumption about orientation in the previous proposition does not present any problem as we can define local orientations. As the loop is contractible (fig 4.18a)) we can find two intervals A, B such that $s(\lambda) \in A$ and $t(\lambda) \in B$ for all $\lambda \in [0,1]$. These two intervals may overlap. Now restrict γ on these two intervals to give curves γ_A, γ_B. Orienting each curve separately so they go the same way round the bitangent circle (as in section 4.4) will give us the situation in the previous lemma. Note if just one of the orientations is reversed then we will find inflections that are on the Symmetry Set and the second part of the Rotational Symmetry Set.

Loops of type 3) pose more of a problem. As in proposition 4.5.1 we look for points on the loop which are also in the set RSS^{-1}. Let M be the value of t for which $\kappa(t)$ assumes its supremum and N the value for which it assumes its infimum, these points are unique for a generic curve. No part of the set RSS^{-1} can cross the lines $M \times I$, $N \times I$, $I \times M$ or $I \times N$ except along the diagonal. From fig 4.20a) we can see that the segment of the $\kappa^{-1}(0)$ contour which starts at M must end at N and from fig 4.20b) we see that any loop which goes round the Mobius band (as



a) In the unit interval. b) On a mobius band.

fig. 4.20) The segment of $K^{-1}(0)$ between the extremum M and N.

in fig. 4.18c)) must cross this segment. Hence loops of type 3) have at least two points for which $\kappa(t_1)$ and $\kappa(t_2)$ are equal.

Note if the bitangent circle is oriented in opposite directions round γ at one of these points then the point does not give an inflection on the SS. The condition for an inflection here is $\kappa(t_1) = -\kappa(t_2)$. It is an open question as to whether such a loop can exist in SS_2^{-1} and if one does wether any point on it will give an inflection on the Symmetry Set.

To find all the closed loops we first calculate both parts of the Rotational Symmetry Set and make a list of all the points where $h(s,t)$ is zero. Then we calculate the parts of the Symmetry Set starting from vertices and the ends of segments, points encountered which are in the list being removed. Finally we start from each of the remaining points in the list. This procedure will find the whole of the Symmetry Set.

Chapter 5: Examples of transitions on the Rotational Symmetry Set, Symmetry Set and Mid-Point Locus

By examining the pictures in this chapter much has been learnt about the various sets and they have helped in the mathematical proof of many results [Tari]. Furthermore they first indicated the connection between the Rotational Symmetry Set and the Dual of the Symmetry Set [Giblin-Tari]. With suitably chosen examples information which could not easily be obtained from mathematical calculation can be found. For example the definition of the Midpoint Locus does not lead to easy mathematical analysis ([Giblin-Brassett] section 5) but examples can be found which clearly illustrate the transitions which occur (fig. 5.42)).

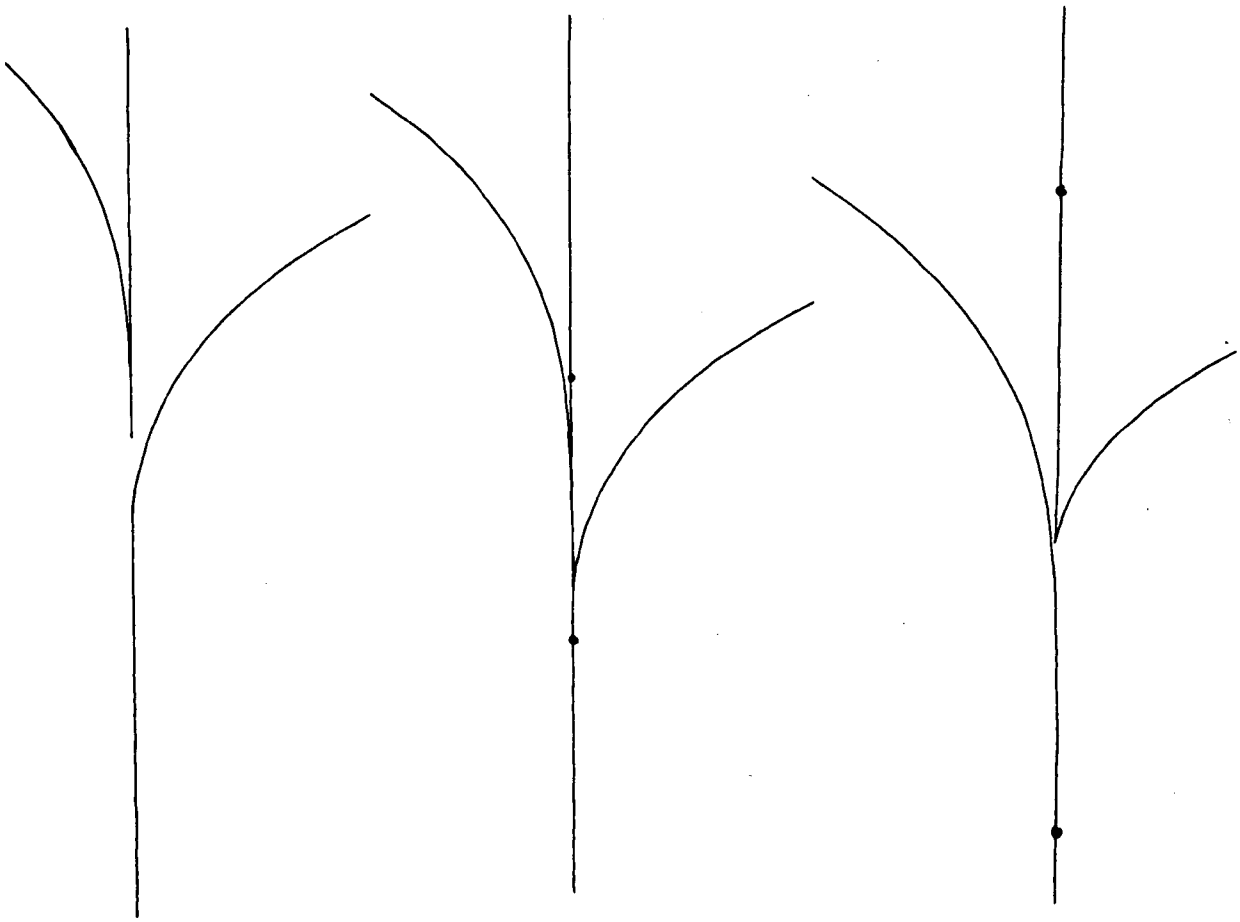


fig. 5.1)* A non-generic transition on the RSS.

Care needs to be taken when choosing examples as misleading results can be obtained. These can be due to choosing a non generic family of curves (the family

of quartic polynomial function always has a higher cusp on its Midpoint Locus [Giblin-Brassett] 5.1,5.4 but for generic curves there are no cusps at all). Even if the family of curves is generic results can be confusing because either two transitions occur almost simultaneously or because a feature is so small that it is not visible or is obscured by the accuracy of the computation. The first attempts at finding the I-beak transition (§5.7) on the RSS were a good example of this (fig. 5.1)*. We found that at the moment of transition there were two points $\gamma(t_1)$, $\gamma(t_2)$ on the curve for which

$$\kappa(t_1) = \kappa(t_2)$$
$$\kappa'(t_1) = \kappa'(t_2) = 0$$
$$\kappa''(t_1) \approx \kappa''(t_2),$$

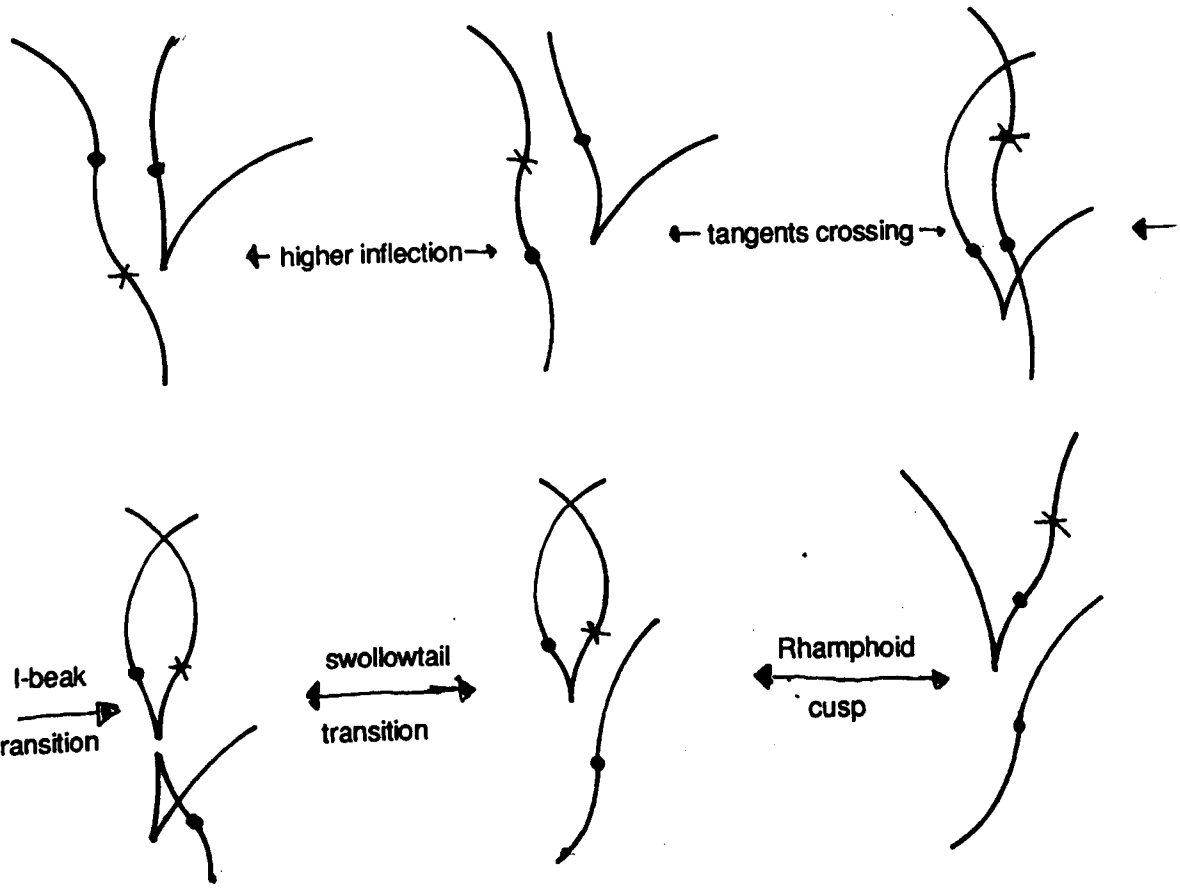


fig. 5.2) Generic transitions which combine to give fig. 5.1).

This has one more condition than would be expected in a one parameter family. Here an I-beak and a Swallowtail transition occur almost simultaneously. In fact the transition was more complex than this as there was a nearby pair of points $\gamma(t_3)$, $\gamma(t_4)$ on the curve for which there was a bitangent circle and the curve had

equal curvature here. This gave rise to an additional inflection on the RSS. The full sequence of transitions which occurred is as shown in figure 5.2).

All the examples in this chapter are of families of curves and their associated transitions on the RSS, SS or MPL. Sections 5.3 to 5.9 only involve the Rotational Symmetry Set and Sections 5.10 to 5.14 involve both the Symmetry Set and the Rotational Symmetry Set. There are four transitions which affect the Midpoint Locus, the two biosculating cases (§5.12 and 5.13) and two instances of cusp transitions (§5.14). Most of the transitions on the RSS were discovered jointly by Farid Tari [Tari] and I. The transitions on the SS and MPL have been known for some time [Bruce-Giblin-1]. This is the first time examples of many of these have been produced.

The conditions for the above transitions are given in terms of the first part of the RSS only. Most of the transitions can also occur on the second part of the RSS and the conditions are summarized in §5.15. There are also two transitions (§5.17 and §5.18) which can only occur on the second part of the RSS.

Section 5.19 concerns the duals of the transitions, the study of these in [Giblin-Tari] and [Tari] was stimulated by the pictures produced here. Section 5.20 deals with the special case of piecewise circular curves.

In the course of the project it was found that useful information about the RSS could be be found by studying the "κ-κ'" diagram (§5.2). This could be used both theoretically to predict the transitions which occurred (for example §5.6) and also as an aid in choosing the examples. An adaptation of this diagram for analysing the second part of the RSS is described in §5.16.

In the text the examples which have been generated by computer are indicated by a * e.g. fig. 5.1)*. For these examples we use the following colour coding.

5.0 Key to Examples

Original Curve	Black
Rotational Symmetry Set	Red or black
Symmetry Set	Blue
Midpoint Locus	Red
Evolute	Green

Special points are marked as follows

Vertices on original curve

Inflections on original curve

Type I inflections on the Rotational Symmetry Set

Type II inflections on the Rotational Symmetry Set

Inflections on the Symmetry Set

§5.1 Generating Examples

Some of the conditions for the various transitions to occur can be quite subtle so a large amount of control over the curves used to generate examples was needed. The types of curves actually used fall into two types:-

Type 1 A single closed curve;

Type 2 Two or more disjoint segments of curve.

For the first type three ellipses were defined in polar form

$$r_i(\theta) = l_i/(1 - e_i\ cos(\theta - \phi_i)\ ),$$

with $i = 1, 2, 3$, $\theta, \phi_i \in [0, 2\pi)$, $l_i > 0$ and $e_i \in (0, 1)$. The curve used was generated

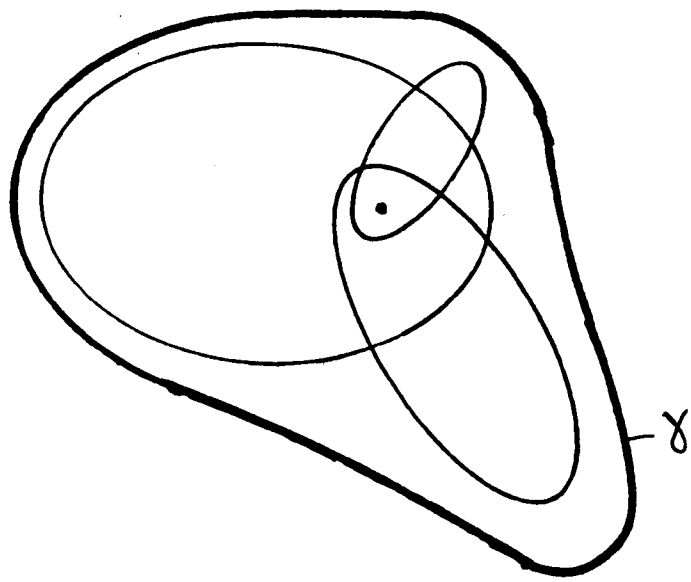


fig. 5.3) A curve produced by taking the sum of three ellipses.

by taking the "sum" of these three ellipses (fig. 5.3), i.e. for $t \in [0, 1)$ we define a curve in polar coordinates

$$\gamma(t) = [R(t), \theta(t)]$$

where

$$\theta(t) = 2\pi t$$

$$R(t) = \sum_{i=1}^{3} r_i(\theta(t)).$$

The shape of this curve is controlled by nine parameters l_i, e_i, ϕ_i $i = 1,2,3$ which gave good control over the position and curvature of the maxima of curvature.

The flexibility of this family of curves is not quite sufficient to generate all the examples needed, so disjoint curves were used for these. One example of the second type of family was to use two cubic functions

$$f_i(t) = a_i t^2 + b_i t^3 \qquad i = 1,2 \qquad t \in [-\frac{1}{2}, \frac{1}{2}]$$

and translate and rotate these curves in the plane to give two curves

$$\gamma_i = (x_i(t), y_i(t)),$$

where

$$x_i(t) = t\ \cos(\phi_i) - f_i(t)\ \sin(\phi_i) + \overline{x}_i,$$
$$y_i(t) = t\ \sin(\phi_i) + f_i(t)\ \cos(\phi_i) + \overline{y}_i.$$

This ten parameter family $\{a_i, b_i, \phi_i, \overline{x}_i, \overline{y}_i \ : \ i = 1,2\}$ is very flexible and was

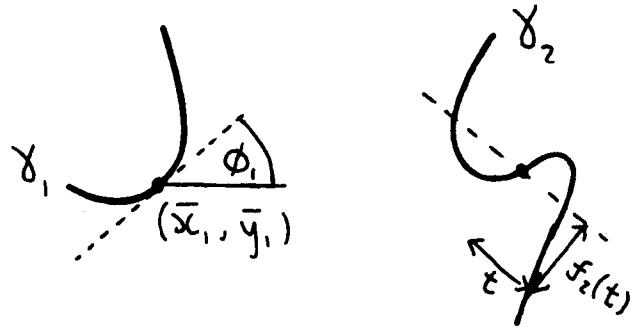


fig. 5.4) Two pieces of cubic curve.

used to generate most of the examples (fig. 5.4). Changing the cubic functions affects the curvature and by changing the rotation and translation of the curves situations where there are bitangent circles can be created or avoided. For example with $\overline{x}_1 = 1$, $\overline{y}_1 = 0$, $\phi_1 = 90°$, $\overline{x}_2 = 0$, $\overline{y}_2 = 1$, $\phi_2 = 180°$ the unit circle centre the origin is tangent to both pieces of curve. We can the alter the cubic function to give the required conditions on the curvature.

For both types of parametrization there is intuitively a high correlation between the parameters used and the actual curve which is produced. This makes both families much easier to use than many other families tried, as special conditions can be engineered rather than having to shoot in the dark to find suitable examples.

Several short programs were written to help choose examples. The most important of these is the κ-κ' diagram. Other methods which were used include looking at the graph of curvature and plotting the signs of $\kappa(s) - \kappa(t)$ or $h(s,t)$ over a course grid in parameter space.

§5.2 The κ-κ' diagram

During the project a useful technique for deducing information about the rotational symmetry set was discovered. This is to look at the κ-κ' graph. This graph is obtained by taking each point on the curve γ and plotting $\kappa'(t)$ (the derivative of curvature with respect to arc length) against $\kappa(t)$ (the curvature). Successive points on this diagram are joined up to produce a closed curve which intersects itself (fig 5.5)*.

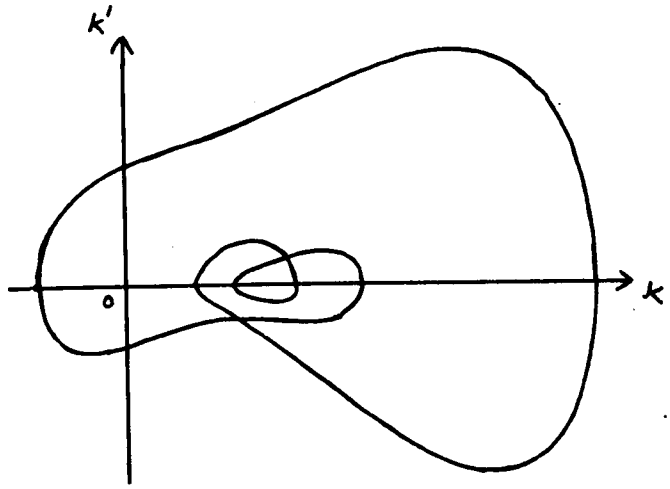


fig. 5.5)* A typical *k-k'* diagram.

The tangent vector to this curve is given by $(\kappa'(t), \kappa''(t))$ so the curve is C^1-differentiable whenever $\kappa' \neq 0$ or $\kappa'' \neq 0$. Hence critical points only occur on the line $\kappa' = 0$. Let t_0 be a point on γ such that $\kappa'(t_0) = \kappa''(t_0) = 0$ the third degree Taylor polynomial for the curvature is given by

$$\kappa(t) = \kappa(t_0) + \tfrac{1}{6}(t - t_0)^3.\kappa'''(t_0) + O(t - t_0)^4.$$

Differentiating gives

$$\kappa' = \tfrac{1}{2}(t - t_0)^2.\kappa'''(t_0) + O(t - t_0)^3$$

so the κ-κ' graph will form a cusp , provided $\kappa'''(t_0) \neq 0$, with a vertical tangent and its critical point lying on the $\kappa' = 0$ line (fig. 5.6).

For a non-degenerate maximum, t_1, or minimum, t_2, of curvature we have $\kappa'(t_1) = \kappa'(t_2) = 0$ and $\kappa''(t_1) < 0 < \kappa''(t_2)$ so the κ-κ' graph always crosses the

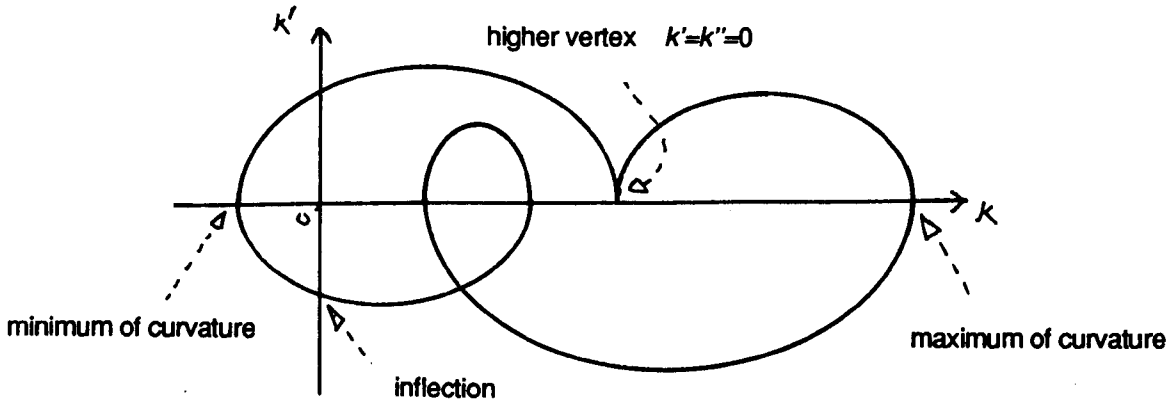


fig. 5.6) The correspondence between points on a curve and its k-k' diagram.

line $\kappa' = 0$ at right angles (fig. 5.6). Furthermore at such a crossing point the curve will always lie on one side of the line $\kappa =$ constant.

Each pair of points on the diagram which have equal values of κ will correspond to a single point on the Rotational Symmetry Set. Following such pairs of points in a continuous manner is equivalent to following pairs of points with equal heights on the curvature graph. The diagram is useful because cusps and type I inflections

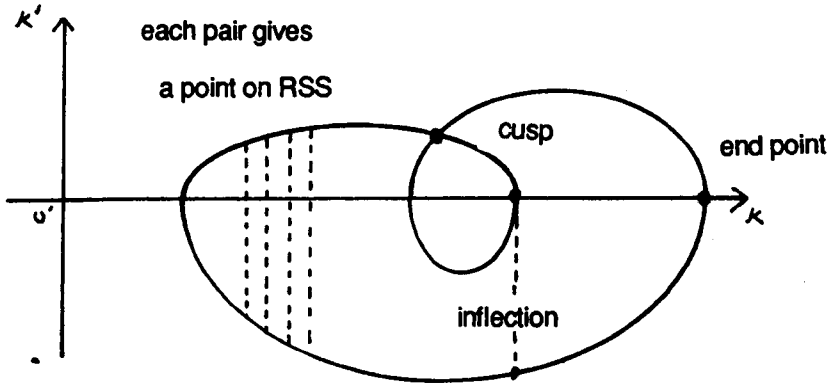


fig. 5.7) The correspondence between the k-k' diagram and the Rotational Symmetry Set.

of the RSS can be detected on it (fig. 5.7). The condition for a cusp is given by $\kappa(t_1) = \kappa(t_2)$ and $\kappa'(t_1) = \kappa'(t_2)$ so corresponds to a self intersection of the diagram. A type I inflection occurs when $\kappa(t_1) = \kappa(t_2)$ and $\kappa'(t_1) = 0$ and $\kappa'(t_2) \neq 0$. On the diagram this corresponds to where one of the branches crosses the $\kappa' = 0$ line. The presence of type II inflections can not be detected from this diagram. By analysing the generic changes which can occur on this diagram we can discover most of the transitions which occur on the RSS, although this in itself does not prove that these transitions are generic for a 1-parameter family of curves.

§5.3 The A_4 Transition on the RSS

In a family of curves γ_t vertices are created in pairs , one maximum and one minimum of curvature. When this happens a new piece of the Rotational Symmetry Set is created. This piece has one cusp, two inflections and two end points. This transition corresponds to an A_4 transition on the original curve.

First we look at how the curvature graph changes in this transition (fig. 5.8). A maximum and minimum are created out of a single point t for which $\kappa'(t) = \kappa''(t) = 0$. From chapter 3 we see that the creation of a new maximum and minimum will

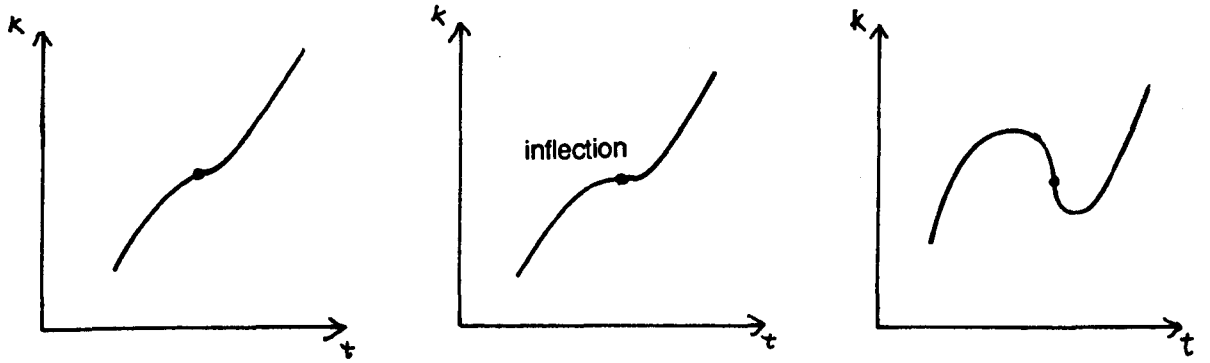


fig. 5.8) A_4 transition:- creation of maxima and minima on curvature graph.

give rise to a new segment of the Rotational Symmetry Set starting and ending at the centres of curvature of the new vertices.

To obtain new information about the shape of this segment we look at the

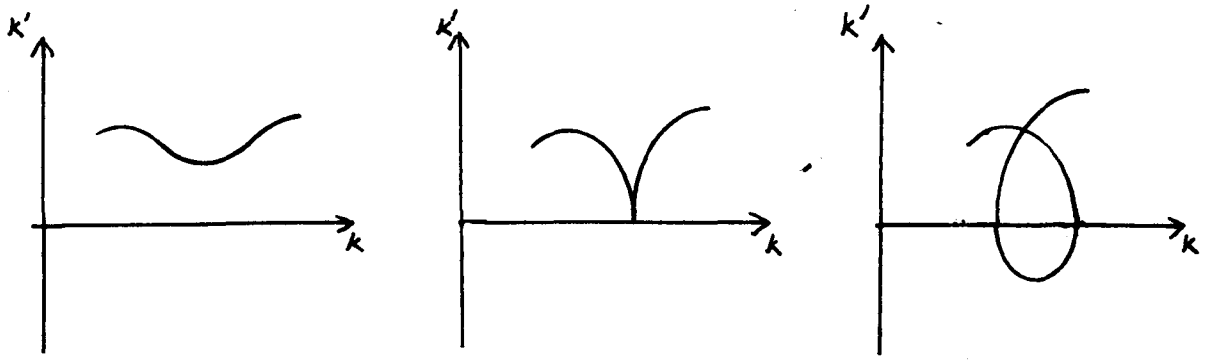


fig. 5.9) A_4 transition:- creation of maxima and minima as seen on k-k' diagram.

$\kappa - \kappa'$ diagram (fig. 5.9). Here part of the curve approaches the $\kappa' = 0$ line and at the moment of transition touches the line in a cusp. After the transition a loop forms which intersects itself.

We see immediately that there will one new cusp and at least four new inflections (two on the new segment). To see how these relate to each other we trace out pairs of points on the diagram with equal values of κ. In (fig. 5.10) the point 1) corresponds to one end of the new section of the RSS. Tracing to the left pairs

2) , 3) are found . One of the points of pair 4) is on the $\kappa' = 0$ line so this pair corresponds to an inflection. We then trace to the right passing through 5) to pair 6) which corresponds to a cusp on the RSS. After the cusp there is another inflection corresponding to pair 8) . Moving left again we end at pair 11) which corresponds to the end of the new section of the RSS. We have now worked out the positions of the cusps and inflection of this new part which looks like figure 5.11), figure 5.35)* shows a computer generated example.

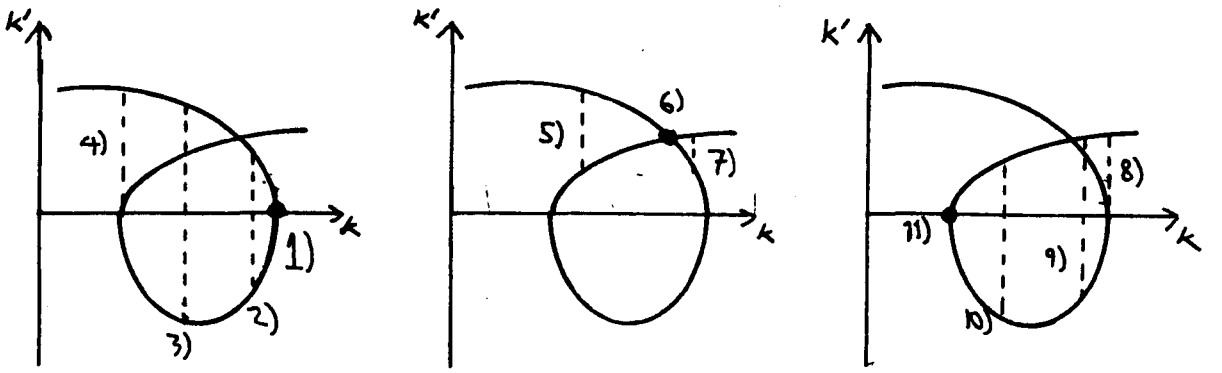


fig. 5.10) A_4 transition:- tracing out pairs of points on k-k' diagram.

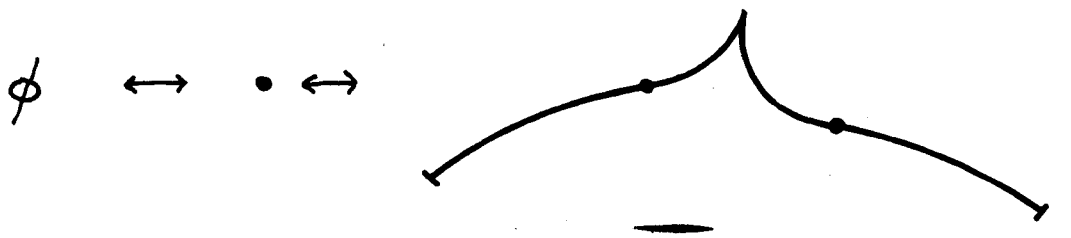


fig. 5.11) A_4 transition:- new section of Rotational Symmetry Set.

As well as a new section pairs of inflections will be created. In figure 5.12) we see that before the transition there were no inflections given by the pairs between 1) and 2). After the transition the pairs 2) , 4) correspond to inflections. Such pairs will be created on each part of the RSS which is given by pairs of points $\gamma(t_1)$, $\gamma(t_2)$ where $\gamma(t_1)$ lies between the two new vertices and for which $\kappa(t_1) = \kappa(t_2)$.

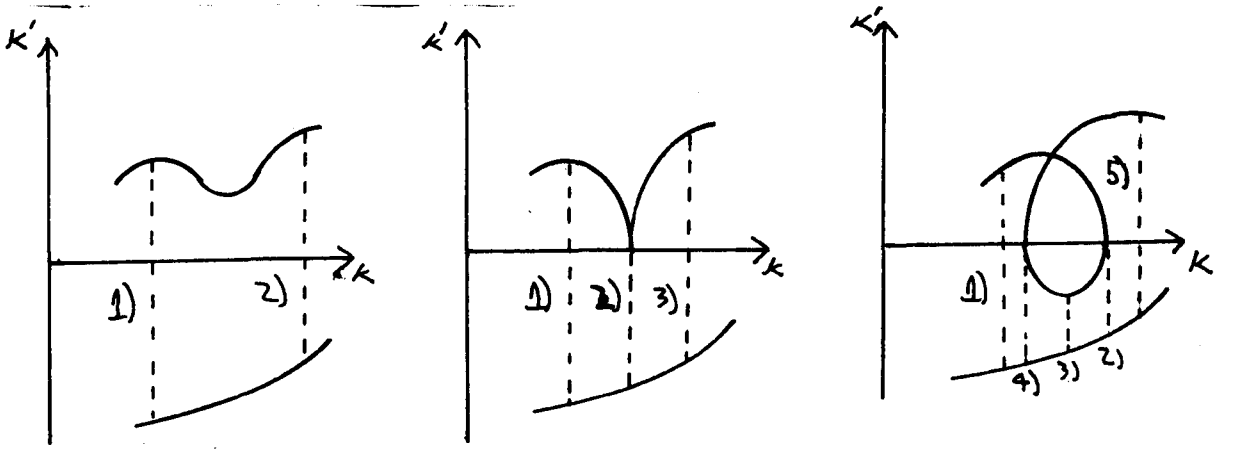


fig. 5.12) A_4 transition:- creation of pairs of inflections as seen on k-k' diagram.

§5.4 The swallowtail transition

When the transition illustrated in figure 5.13) occurs on the κ-κ' diagram and the point of contact x does not lie on the $\kappa' = 0$ line we see that two cusps (1) and

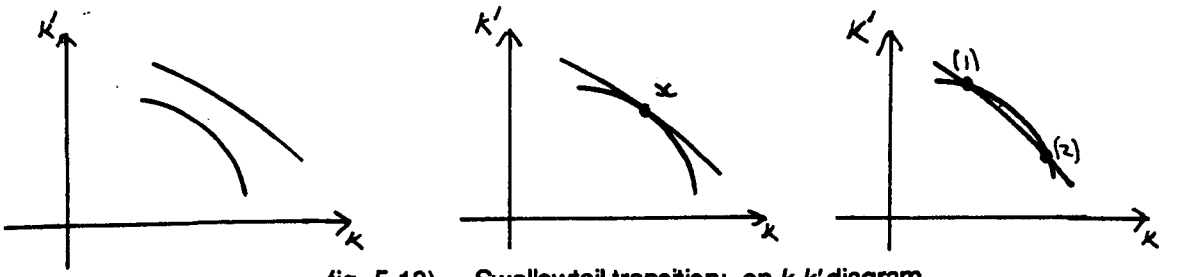


fig. 5.13) Swallowtail transition:- on k-k' diagram.

(2) are created and generically there will not be any inflections created. This will produce a swallowtail transition on the RSS (fig. 5.14). For the transition to occur there must be two points $t_1, t_2 \in \gamma$ such that $\kappa(t_1) = \kappa(t_2)$, $\kappa'(t_1) = \kappa'(t_2) \neq 0$ and the two parts of the κ-κ' diagram meet tangentially. So we have the extra condition $\kappa''(t_1) = \kappa''(t_2)$.

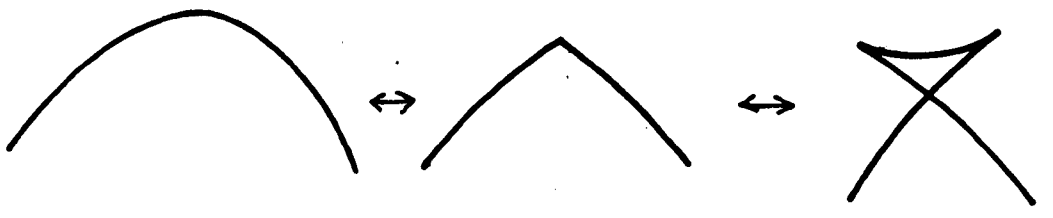


fig. 5.14) Swallowtail transition:- on Rotational Symmetry Set.

§5.5 The I-Lips Transition.

From §3.5 we saw that closed loops could appear in the RSS. The curvature graph of a transition which causes such loops to be formed is illustrated in figure

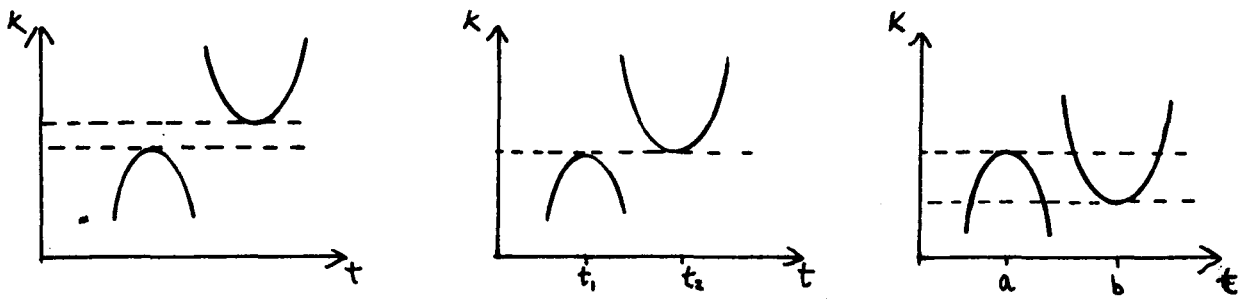


fig. 5.15) l-lips transition:- curvature graph.

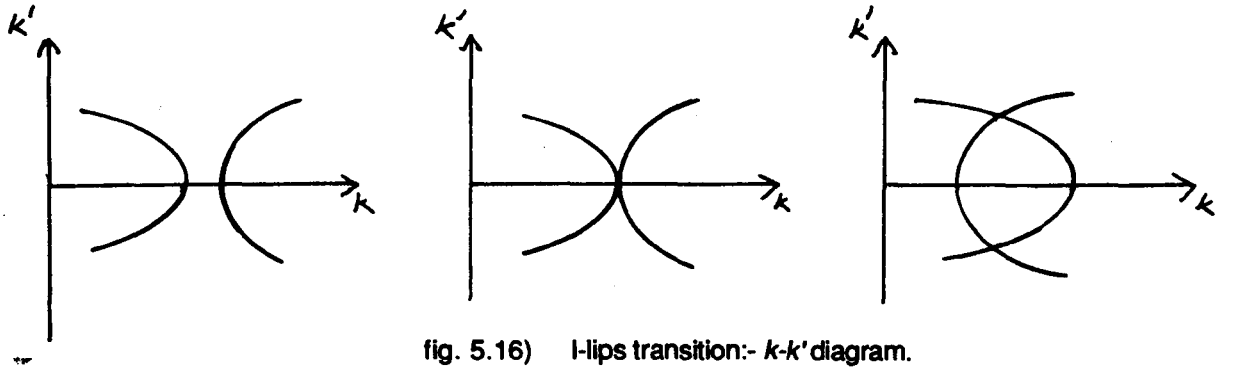


fig. 5.16) l-lips transition:- *k-k'* diagram.

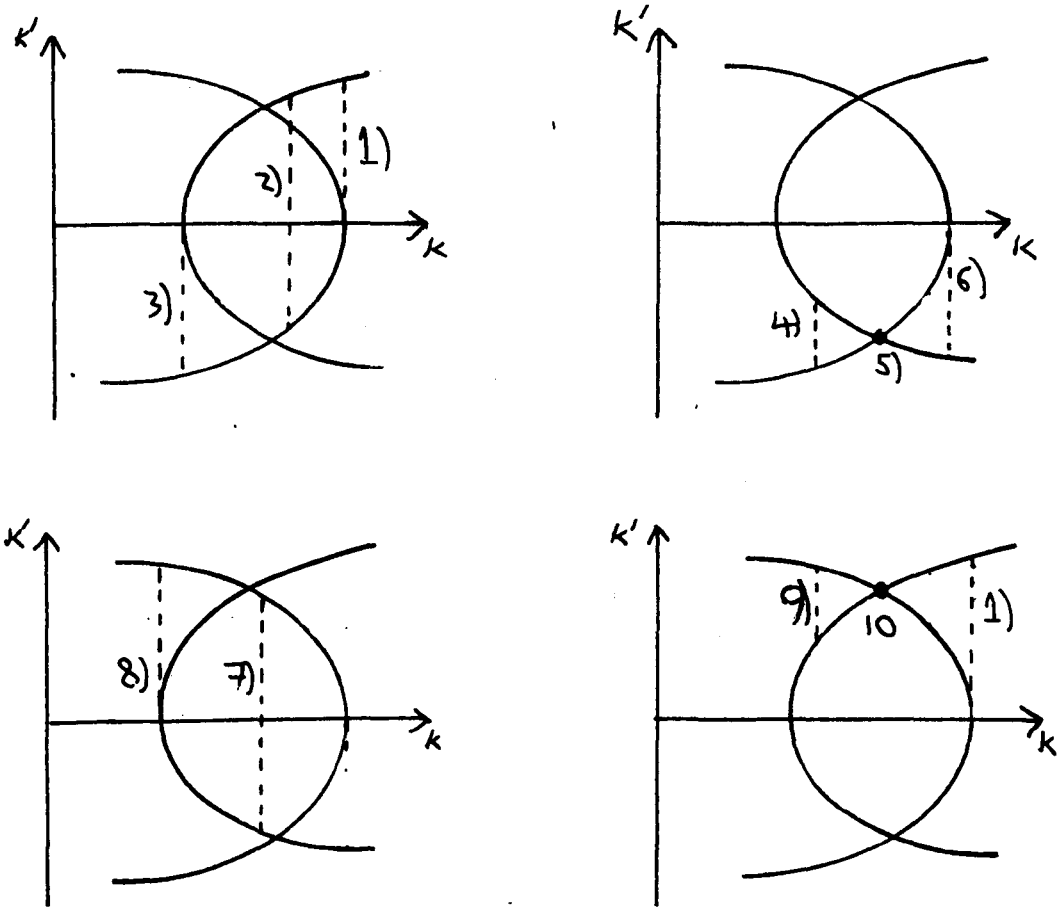


fig. 5.17) l-lips transition:- tracing out pairs of points on *k-k'* diagram.

5.15). After the transition we see that the maximum a and minimum, b, satisfy the "closed loop" condition of §3.5. Before the transition we observe there is no such loop and at the moment of transition there will be an isolated point. The condition for this transition is $\kappa(t_1) = \kappa(t_2)$, $\kappa'(t_1) = \kappa'(t_2) = 0$ and $\kappa''(t_1) < 0 < \kappa''(t_2)$. To find the structure of this loop we look at the κ-κ' diagram (fig. 5.16). We now trace out pairs of points on the κ-κ' graph with equal curvatures . As we have a closed loop we can start tracing from any suitable pair. We will start at the pair 1) of (fig. 5.17) which gives an inflection on the RSS. Moving to the right we pass through 2), 3) – Inflection, 4), 5) – Cusp, 6) – Inflection, 7), 8) – Inflection, 9), 10) – Cusp and finally back to 1). This gives the transition on the RSS illustrated in

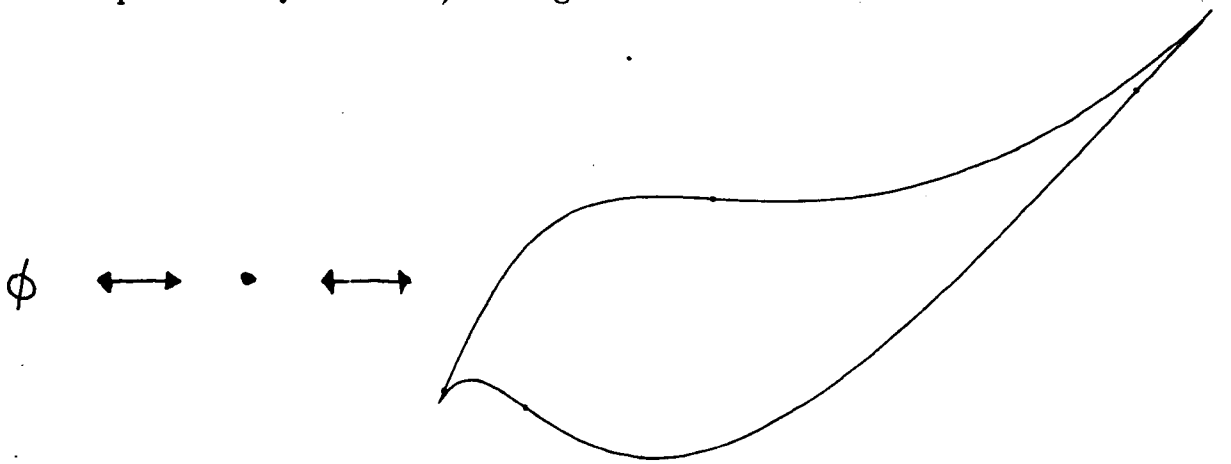


fig. 5.18)* I-lips transition:- on Rotation Symmetry Set.

figure 5.18)*. This is similar to the normal "Lips" transition except there are four inflections instead of two . From studying the dual of this transition Farid Tari has shown there is also a bitangent line [Tari]. We will call this the **I - Lips transition** to indicate the presence of the extra inflections.

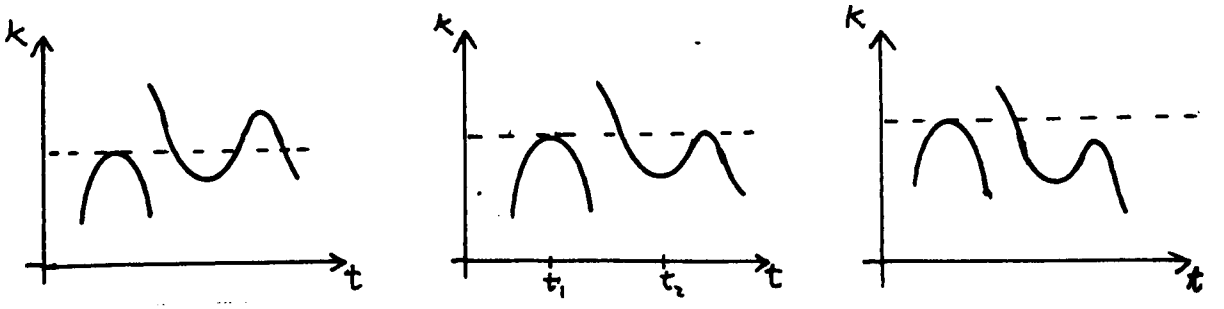


fig. 5.19) One way an I-lips can break up:- curvature graph.

There are a few ways in which this closed loop could break up which correspond to one of the conditions $2a') \ldots 2d')$ of §3.5 failing. For example we could have the

curvature of the maximum t_1 increasing until it was greater than that of another maximum of curvature. On the curvature graph it looks like figure 5.19). This would generically correspond to an $I-beaks$ transition (§5.7) which would join the loop to an other part of the curve as illustrated in figure 5.20).

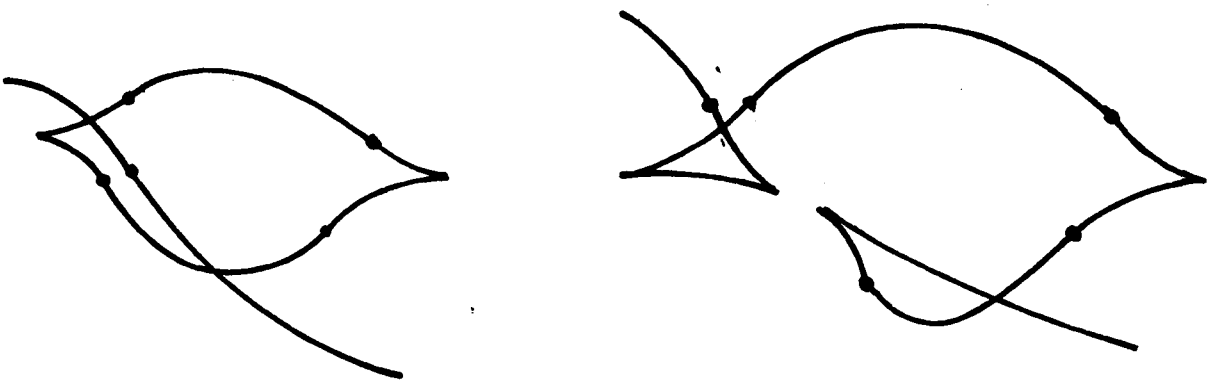


fig. 5.20) One way an I-lips can break up:- Rotational Symmetry Set.

§5.6 Two equal maxima of curvature

These two transitions occur when two vertices (both maxima or both minima) have equal values of curvature, as shown in (fig. 5.21). When we look at the κ-κ'

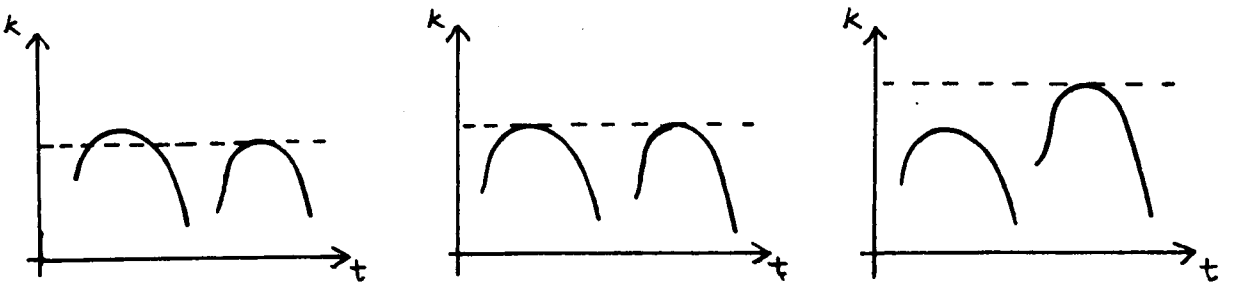


fig. 5.21) Equal maxima of curvature:- curvature graph.

diagram we see this transition can happen in one of two ways (fig 5.22). Let $\gamma(t_0)$, $\gamma(t_1)$ be the two vertices. To distinguish between the cases we look at the functions

$$f(x) = \kappa(t) - \kappa(t_0),$$
$$g(x) = \kappa(t) - \kappa(t_1),$$

where $x = \kappa'(t)$ (fig. 5.23). We will only consider the generic situation where $\kappa''(t_0)$ and $\kappa''(t_1)$ are both non zero. In both cases we have $f(0) = g(0) = f'(0) = g'(0) = 0$. The third degree Taylor polynomial approximations for the two functions are

$$f(x) = \frac{1}{2}f''(0)x^2 + \frac{1}{6}f'''(0)x^3 + O(x^4),$$
$$g(x) = \frac{1}{2}g''(0)x^2 + \frac{1}{6}g'''(0)x^3 + O(x^4).$$

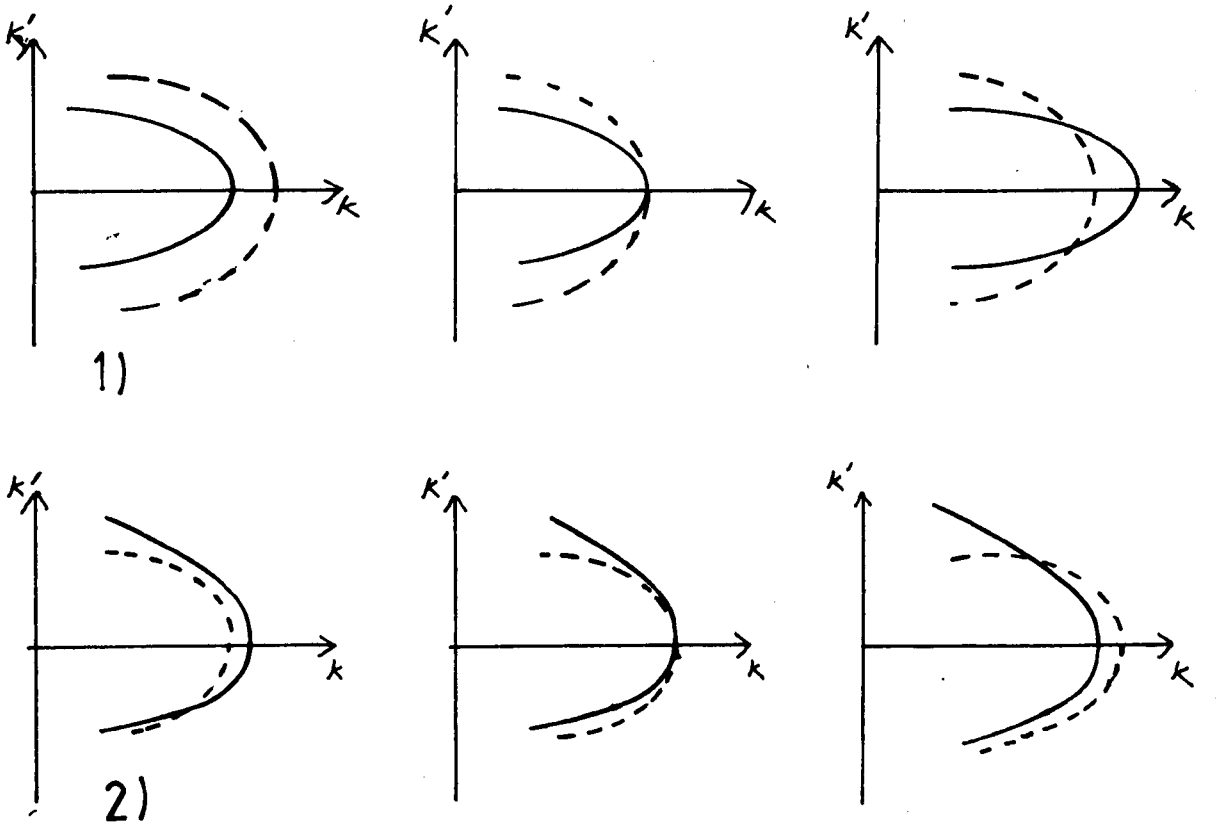


fig. 5.22) Equal maxima of curvature:- two different transitions on *k-k'* diagram.

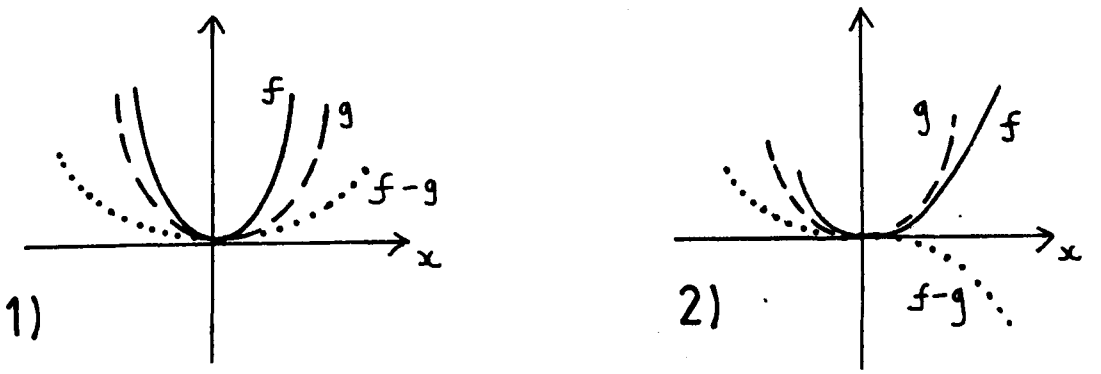


fig. 5.23) The functions f, g, f-g for the two different transitions.

Subtracting gives

$$f(x) - g(x) = \frac{1}{2}(f''(0) - g''(0))x^2 + \frac{1}{6}(f'''(0) - g'''(0)x^3 + O(x^4).$$

If $f''(0) \neq g''(0)$ then $f(x) - g(x) = \alpha x^2 + O(x^3)$ which gives the first transition. If $f''(0) = g''(0)$ and $f'''(0) \neq g'''(0)$ then $f(x) - g(x) = \beta x^3 + O(x^4)$ which is the second situation.

We now find the derivative of the function f in terms of the curvature

$$\begin{aligned}
\frac{df}{dx} &= \frac{d\kappa}{dt}\frac{dt}{dx} = \frac{d\kappa}{dt}\left(\frac{d\kappa'}{dt}\right)^{-1} \\
&= \frac{\kappa'(t)}{\kappa''(t)}. \\
\frac{d^2f}{dx^2} &= \frac{(\kappa''(t))^2 - \kappa'(t)\,\kappa''(t)}{(\kappa''(t))^2}\,\frac{1}{\kappa''(t)}, \\
\frac{d^3f}{dx^3} &= \frac{-2(\kappa''(t))^2\kappa''(t) - \kappa'(t)\kappa'(t)\kappa^{(4)}(t) - 3\kappa'(t)(\kappa'''(t))^2}{(\kappa''(t))^5}.
\end{aligned}$$

Evaluating at the origin

$$\begin{aligned}
\frac{df}{dx}(0) &= \frac{\kappa'(t_0)}{\kappa''(t_0)} = 0, \\
\frac{d^2f}{dx^2}(0) &= \frac{1}{\kappa''(t_0)}, \\
\frac{d^3f}{dx^3}(0) &= \frac{-2\kappa'''(t_0)}{\kappa''(t_0)^3}.
\end{aligned}$$

We now have some conditions for the two transitions

Transition 1 (I - beak transition)

$$\begin{aligned}
\kappa(t_0) &= \kappa(t_1), \\
\kappa'(t_0) &= \kappa'(t_1) = 0, \\
\kappa''(t_0) &\neq \kappa''(t_1).
\end{aligned}$$

Transition 2 (Trident transition)

$$\begin{aligned}
\kappa(t_0) &= \kappa(t_1), \\
\kappa'(t_0) &= \kappa'(t_1) = 0, \\
\kappa''(t_0) &= \kappa''(t_1) \neq 0, \\
\kappa'''(t_0) &\neq \kappa'''(t_1).
\end{aligned}$$

We observe that transition 1 will generically occur in a one parameter family but transition 2 will not. We now examine the κ-κ' graph to find out what transitions occur in the Rotational Symmetry Set for these two situations.

§5.7 I - Beak Transition

We can now see immediately that before the transition there are only two inflections and afterwards there are two inflections and two cusps. We trace out the pairs of points to see how these are arranged. Starting at a in figure 5.24) we pass through 1), 2) – inflection, 3) and end at b. On the other part of the RSS we start at d pass through 5), 6) – inflection, 7) and end at pair c. After the transition (fig. 5.25) we start at pair a' pass through 1'), 2') – cusp, 3') – inflection 4') and end at pair c'. Starting at d' we pass through 5'), 6') – Cusp, 7') – Inflection, 8') and end at pair b' .

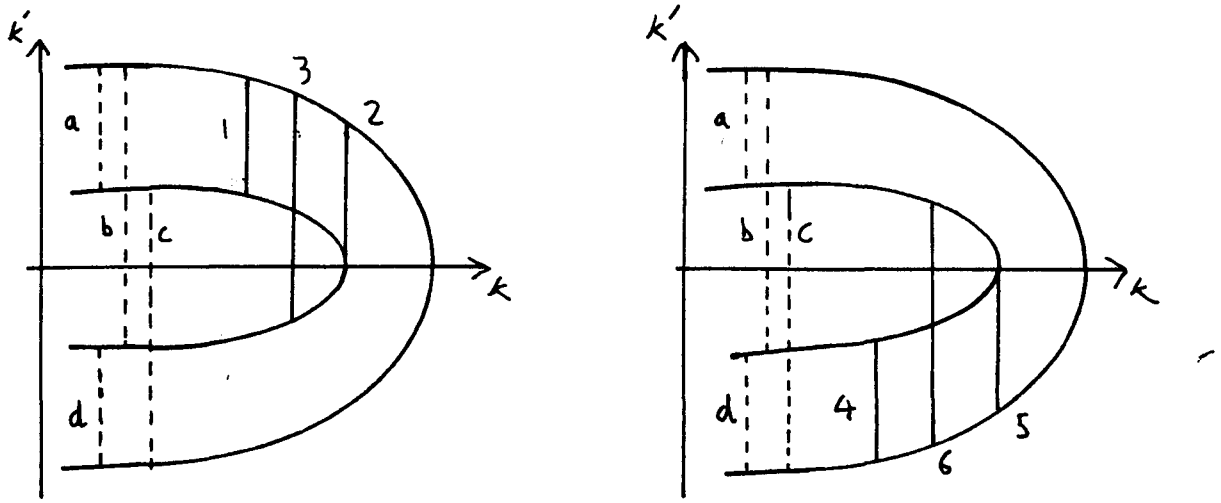


fig. 5.24) I-beak transition:- tracing out pairs of points, before transition.

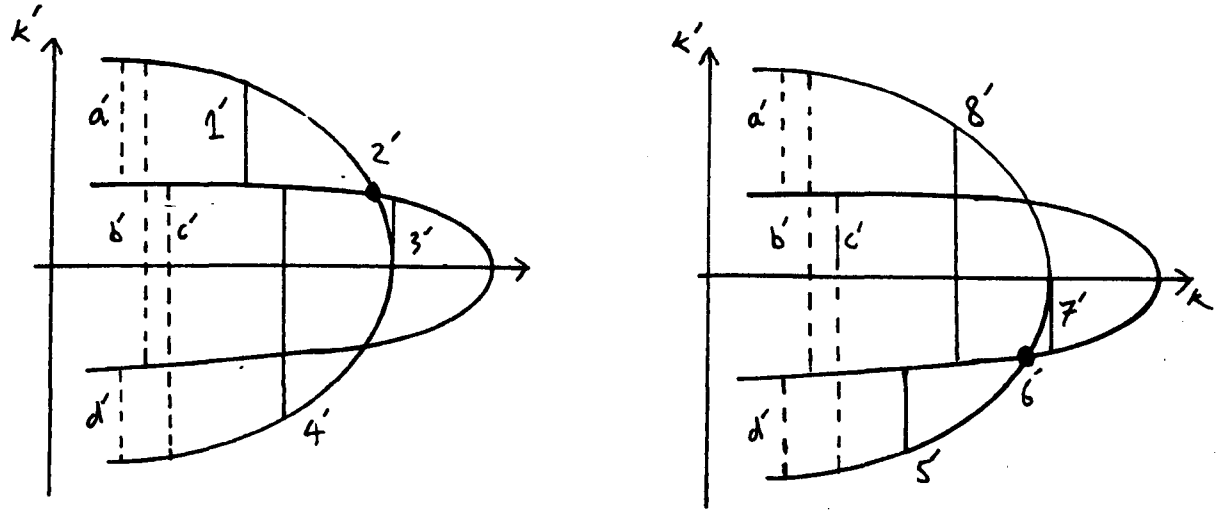
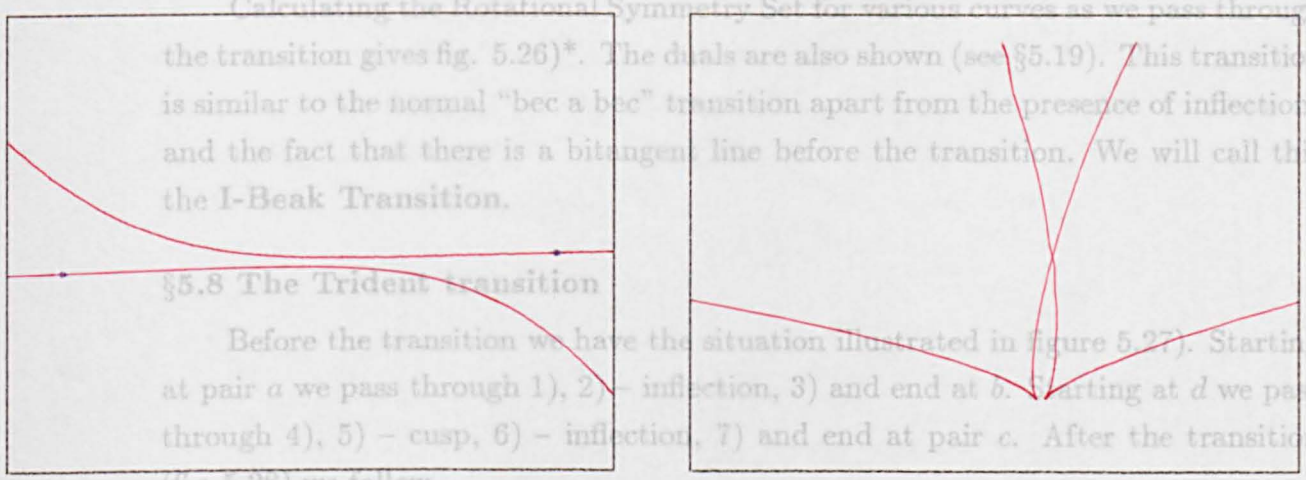
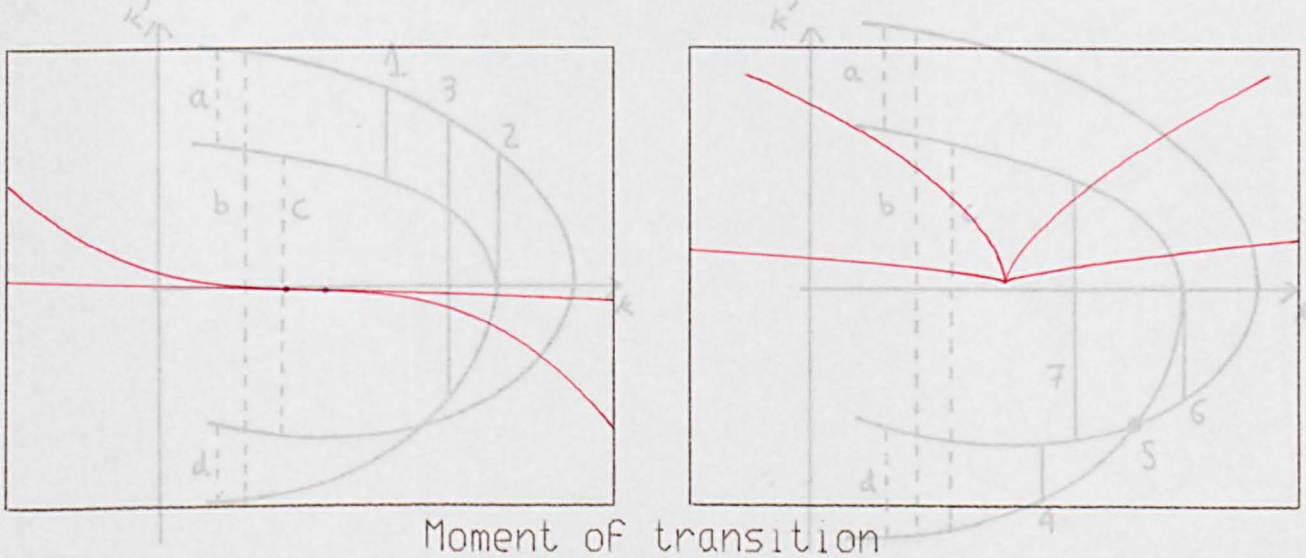


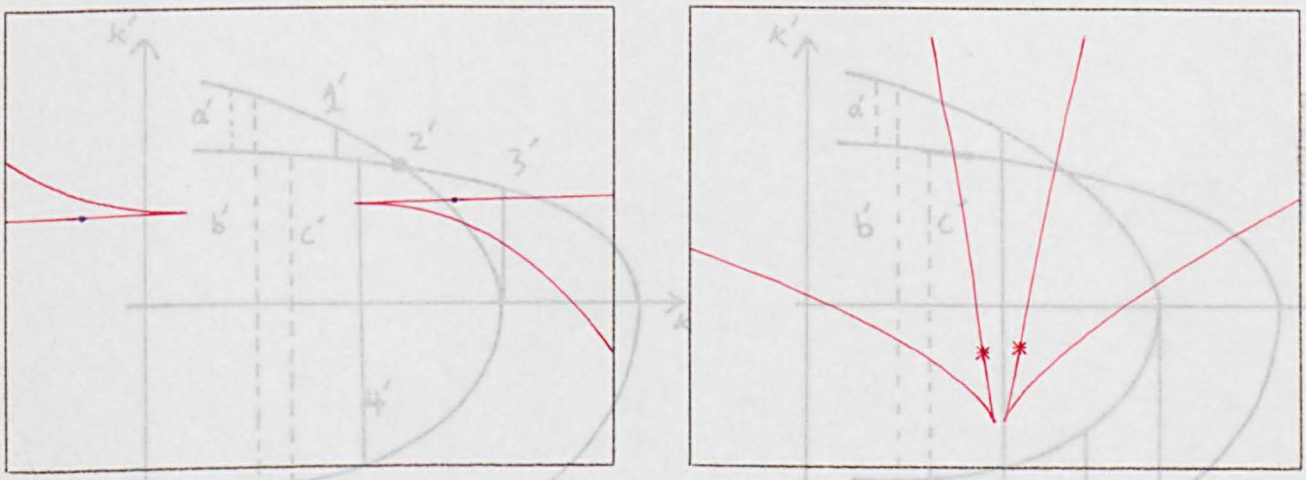
fig. 5.25) I-beak transition:- tracing out pairs of points, after transition.





Moment of transition

fig. 5.27) Trident transition:- tracing out pairs of points, before transition.



Rotational Symmetry Set

Dual

fig. 5.28) Trident transition:- tracing out pairs of points, after transition.

Fig 5.26: I-beak transition on RSS and dual

Calculating the Rotational Symmetry Set for various curves as we pass through the transition gives fig. 5.26)*. The duals are also shown (see §5.19). This transition is similar to the normal "bec a bec" transition apart from the presence of inflections and the fact that there is a bitangent line before the transition. We will call this the **I-Beak Transition.**

§5.8 The Trident transition

Before the transition we have the situation illustrated in figure 5.27). Starting at pair a we pass through 1), 2) – inflection, 3) and end at b. Starting at d we pass through 4), 5) – cusp, 6) – inflection, 7) and end at pair c. After the transition (fig. 5.28) we follow

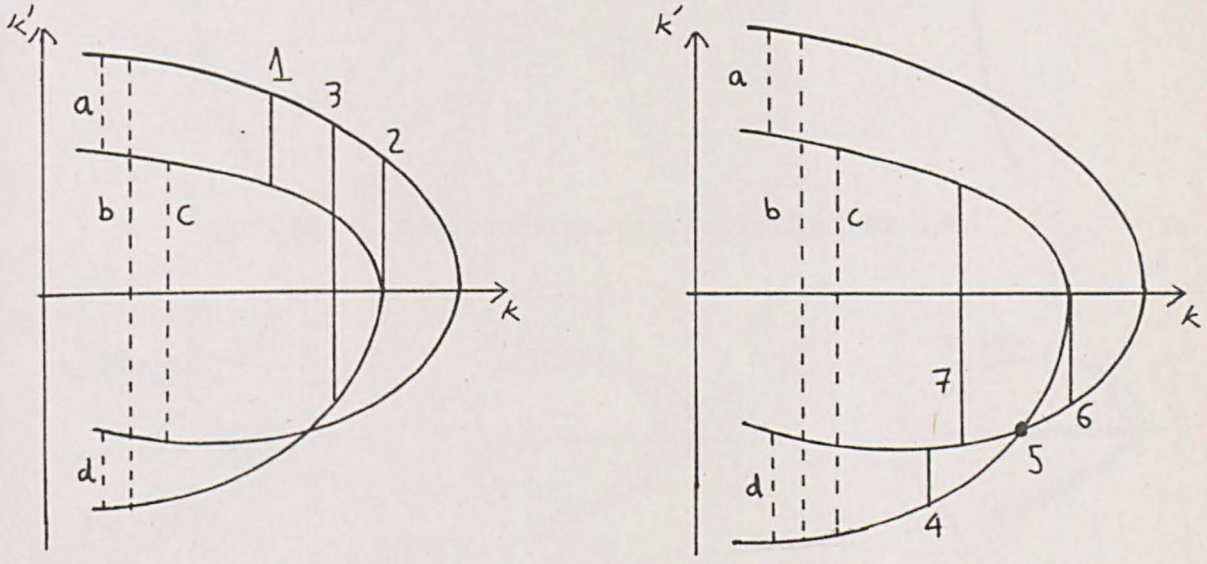


fig. 5.27) Trident transition:- tracing out pairs of points, before transition.

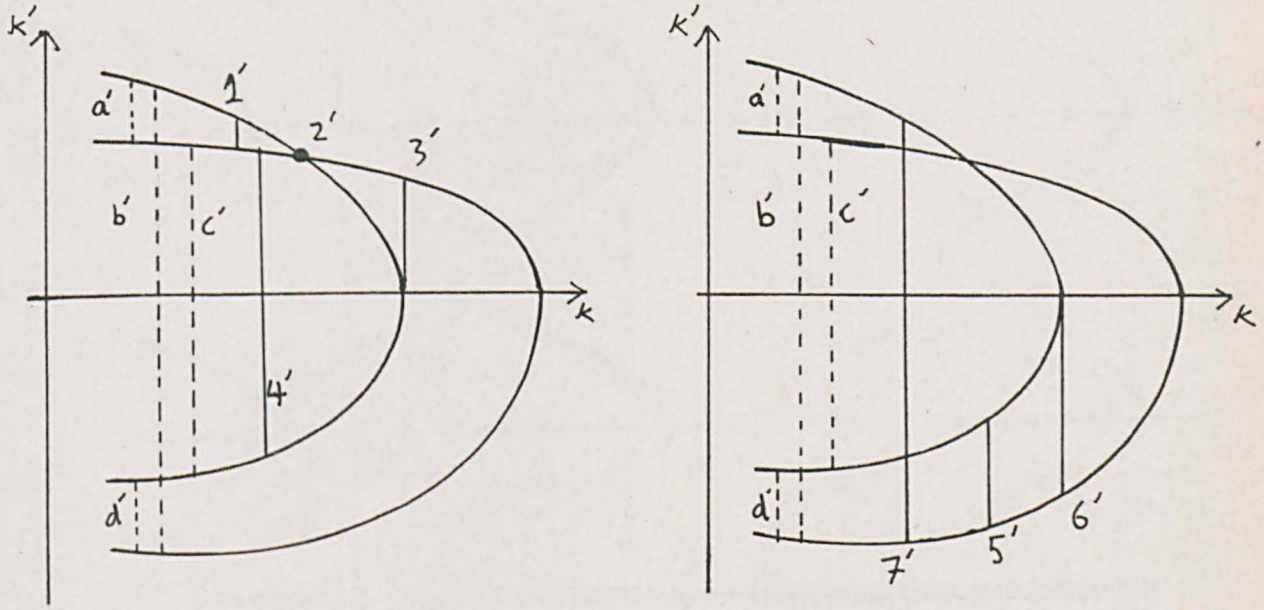


fig. 5.28) Trident transition:- tracing out pairs of points, after transition.

the pairs starting at a' to give the following sequence 1'), 2') – cusp, 3') – inflection, 4') which ends at c'. Now if we start at d' we pass through 5'), 6') – inflection, 7') and end at b' . As the moment of transition is approached the cusp and inflections are brought

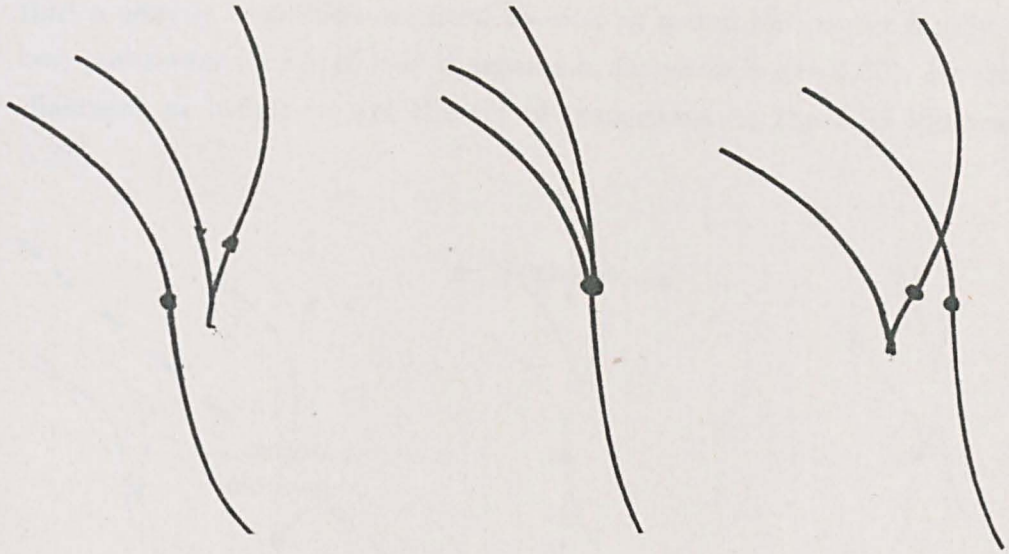


fig. 5.29) Trident transition:- on Rotational Symmetry Set.

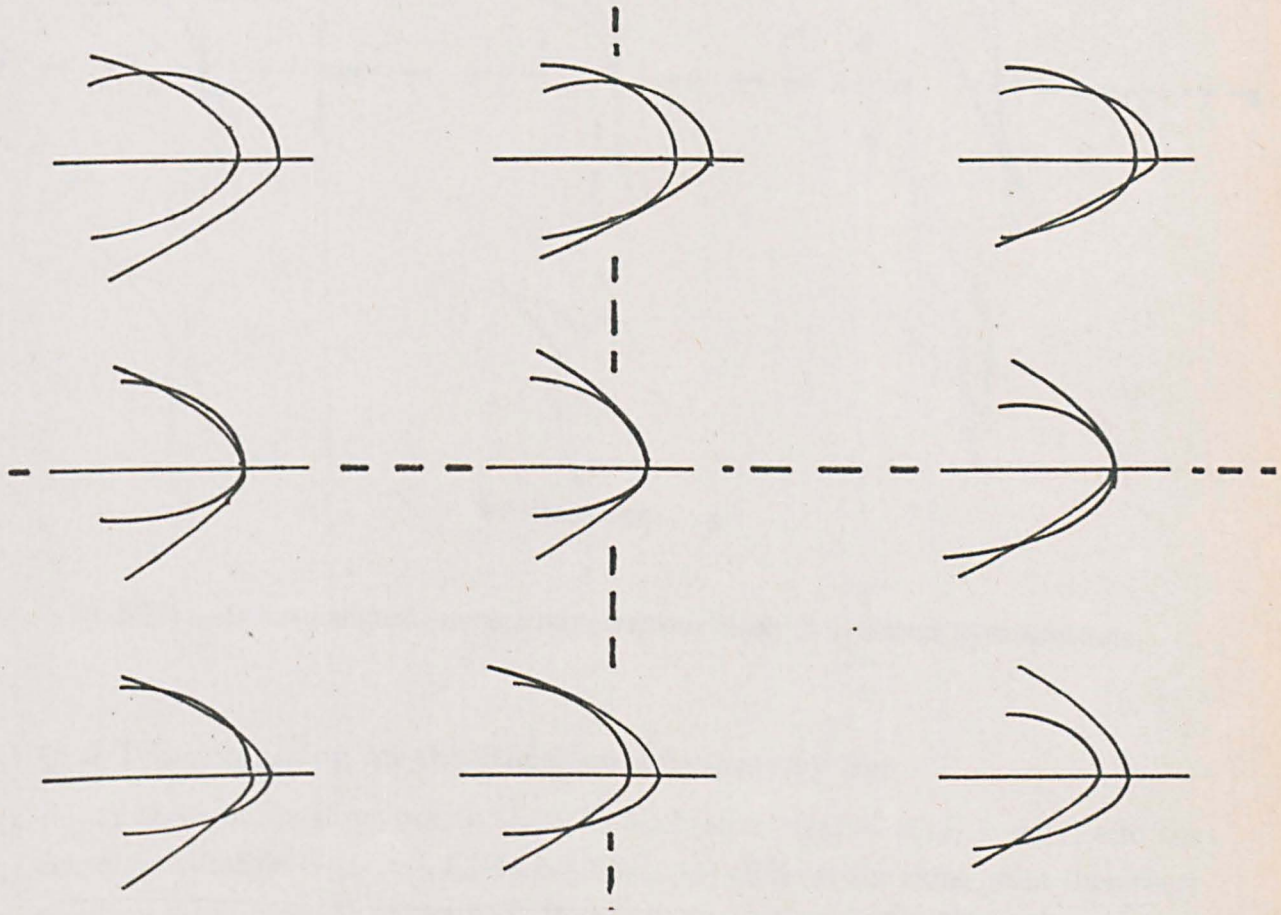


fig. 5.30) Trident transition:- generic two parameter family of *k*-*k'* diagrams.

together into a single point. This point breaks up after the transition into a different arrangement of cusps and inflections but still preserving the number of each. Figure 5.29) indicates the transition which appears on the RSS.

As noted above this transition is not generic in a one parameter family so to find a generic transition we need to look at a two parameter family of curves. A two parameter family of κ-κ' diagrams is shown in figure 5.30). by analysing these diagrams as before we get the set of transitions on the RSS illustrated in figure 5.31).

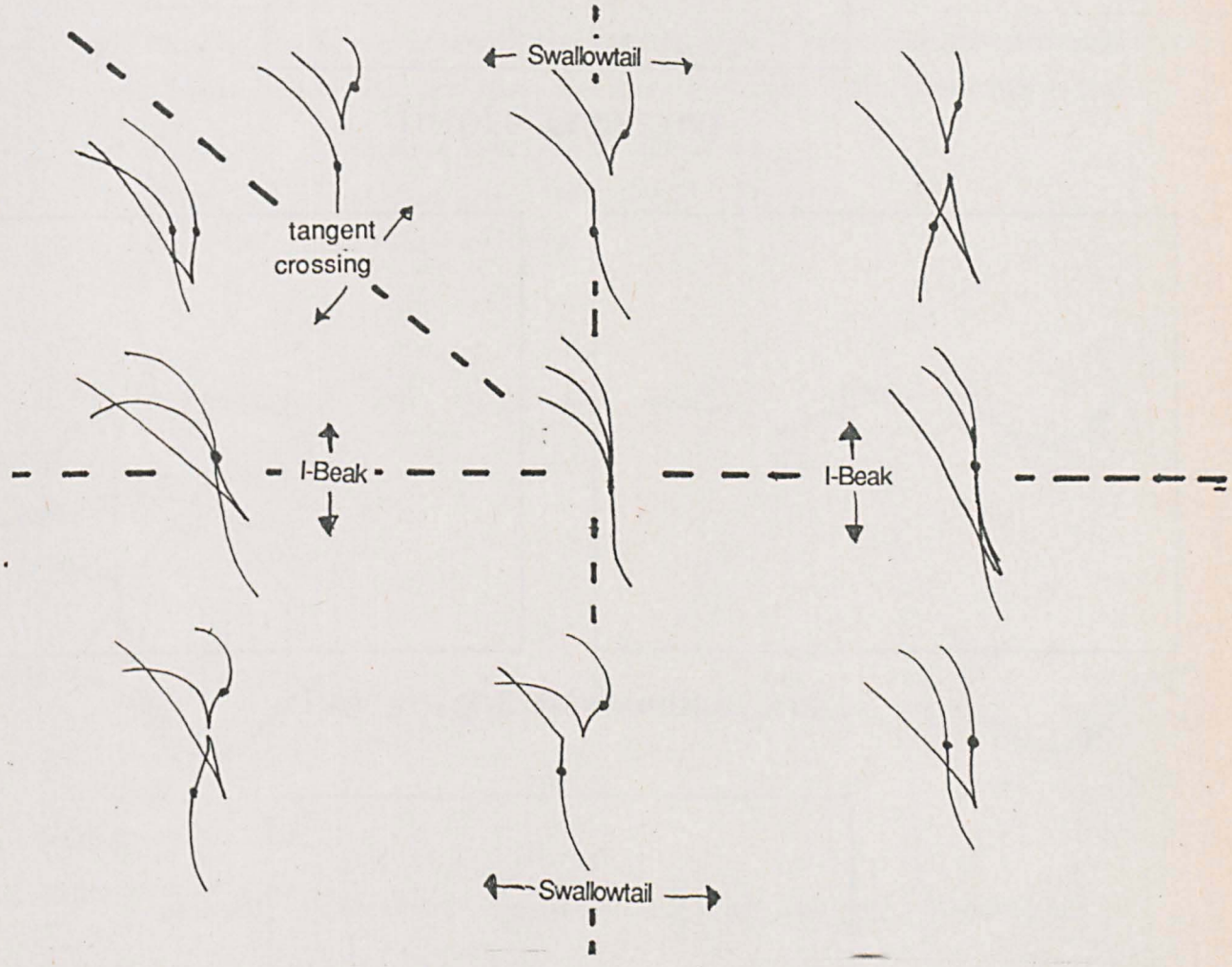
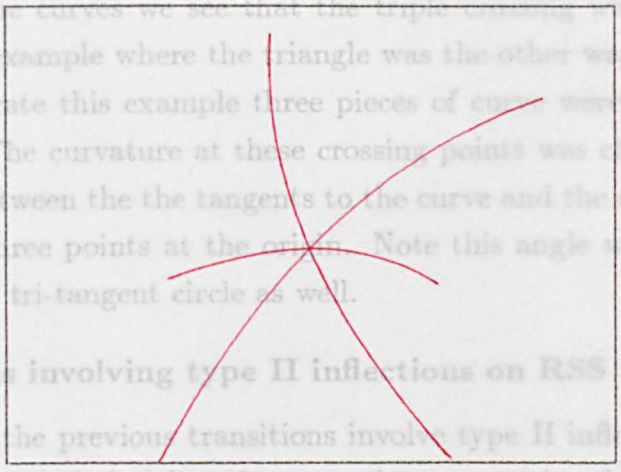


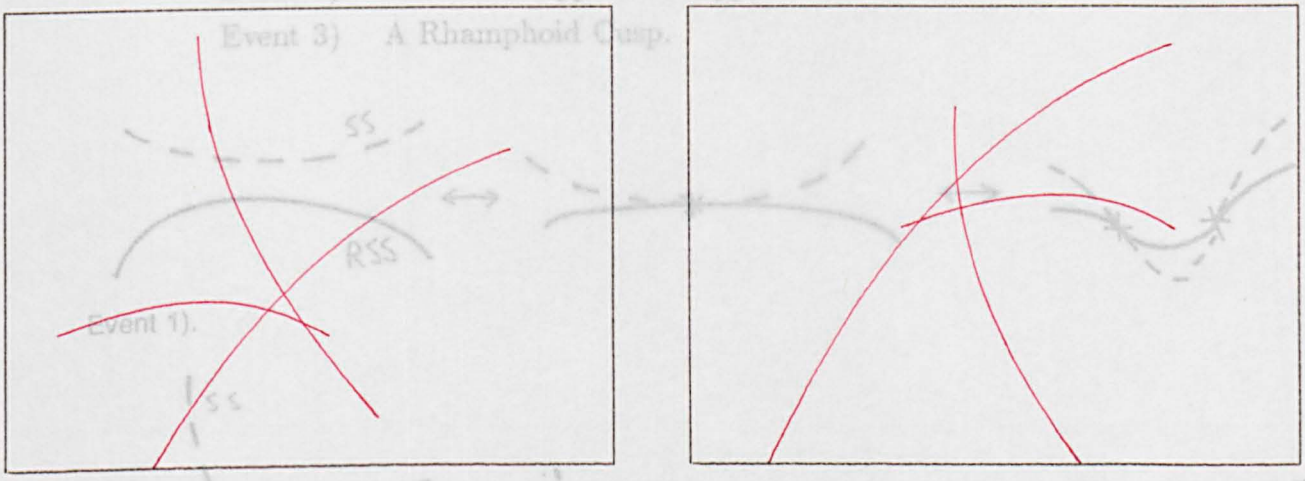
fig. 5.31) Trident transition:- generic two parameter family of Rotational Symmetry Sets.

§5.9 Triple crossing on the Rotational Symmetry Set

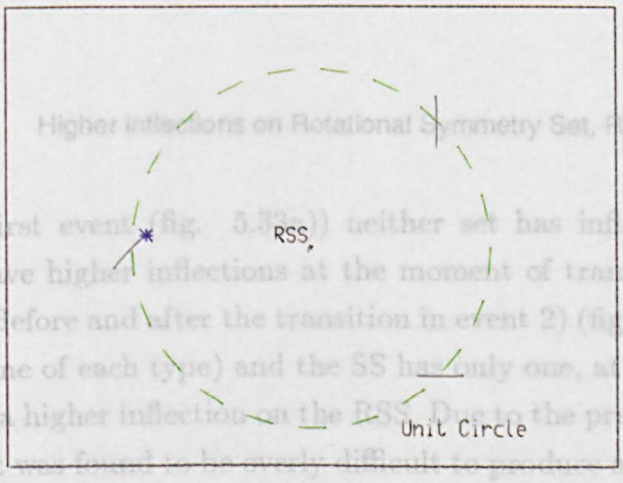
If there exists three points t_1, t_2, t_3 such that $\kappa(t_1) = \kappa(t_2) = \kappa(t_3)$ and the centers of rotation $C_1(t_1, t_2)$, $C_1(t_1, t_3)$, $C_1(t_2, t_3)$ all lie at the same point then there will be a triple crossing of the RSS. By generating an example of this situation and



Triple Crossing



Two slight perturbations



Curves used to generate example

Fig 5.32: Triple crossing on RSS

perturbing the curves we see that the triple crossing was not stable (fig 3.32)*). However an example where the triangle was the other way up was not found.

To generate this example three pieces of curve were used which all cross the unit circle. The curvature at these crossing points was chosen to be the same and the angles between the the tangents to the curve and the unit circle were also equal to produce three points at the origin. Note this angle was not zero otherwise we would have a tri-tangent circle as well.

§5.10 Events involving type II inflections on RSS and SS

None of the previous transitions involve type II inflections (those which arise due to bitangent circles) but there are three events in which these play a part.

Event 1) Creation of two type II inflections.
Event 2) Coincident type I and type II inflections
Event 3) A Rhamphoid Cusp.

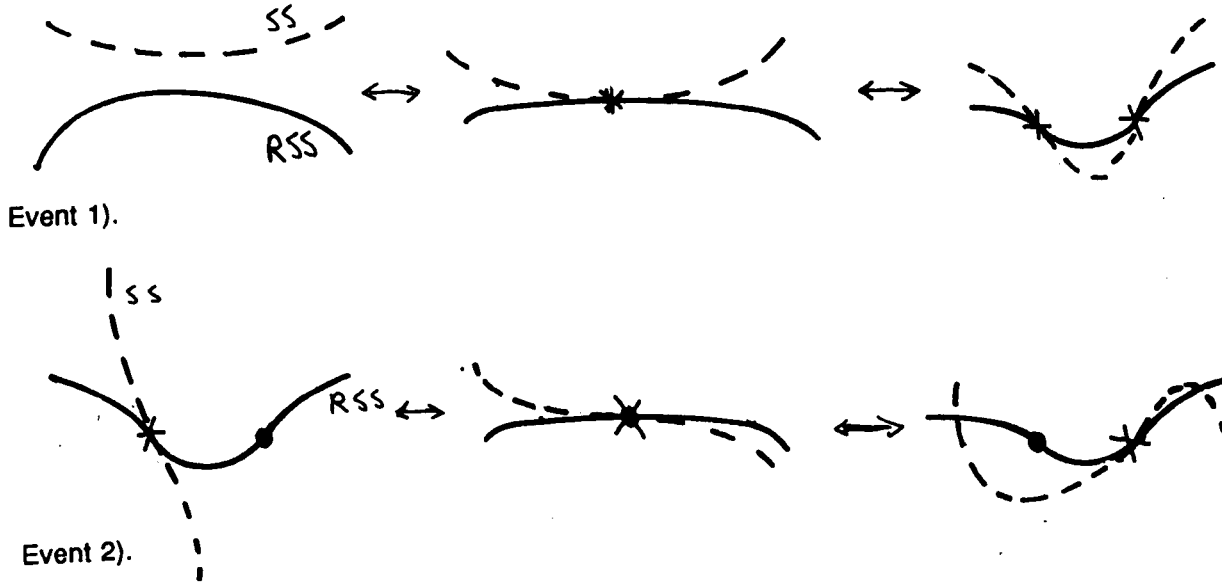


fig. 5.33) Higher inflections on Rotational Symmetry Set, RSS + Symmetry Set.

In the first event (fig. 5.33a)) neither set has inflections before transition, they both have higher inflections at the moment of transition and two inflections afterwards. Before and after the transition in event 2) (fig. 5.33b)) the RSS has two inflections (one of each type) and the SS has only one, at the moment of transition there will be a higher inflection on the RSS. Due to the presence of higher inflections on the RSS it was found to be overly difficult to produce a computer example where the RSS was actually distinguishable from a straight line.

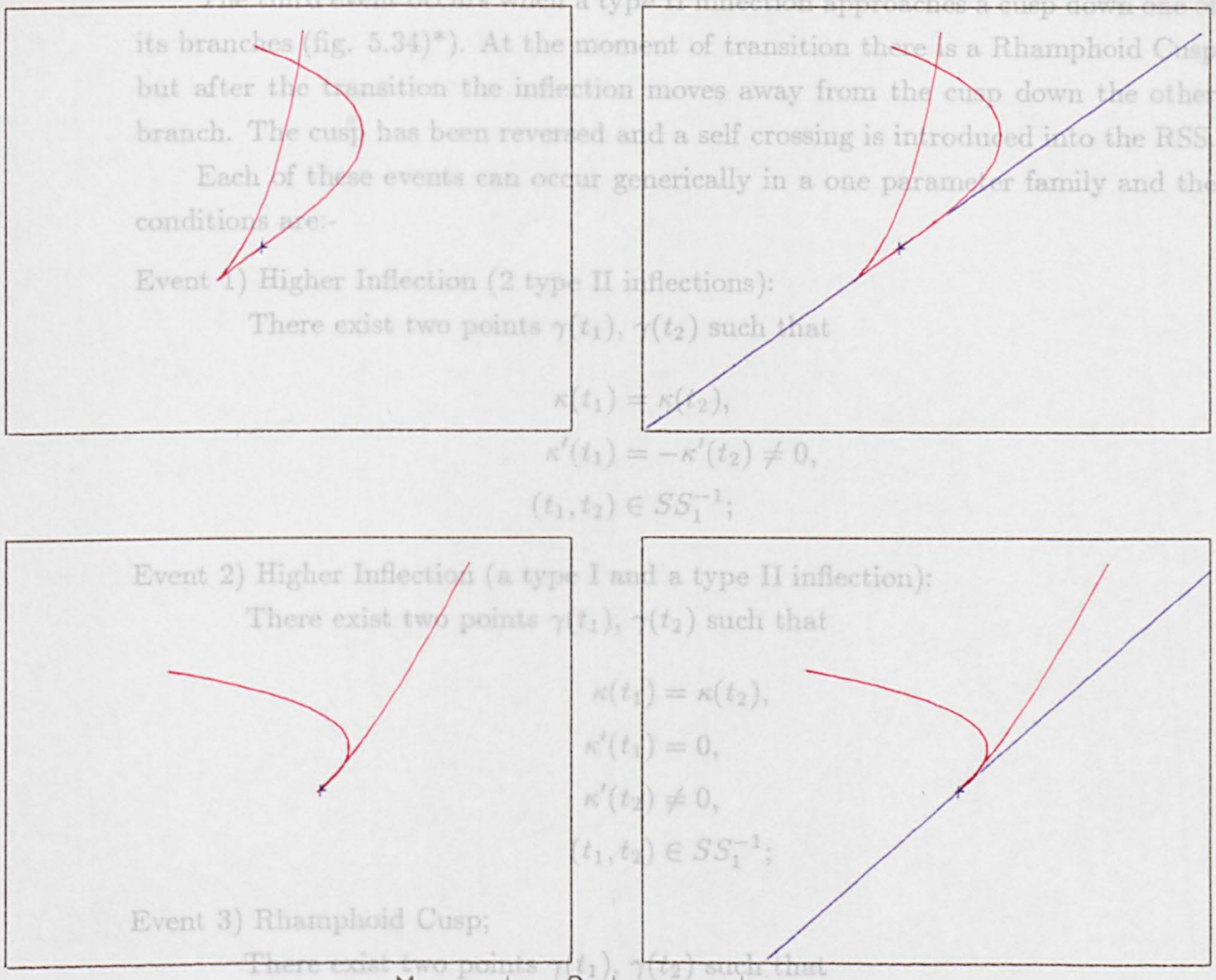


Fig 5.34) Rhamphoid Cusp on RSS

The third event occurs when a type II inflection approaches a cusp down one of its branches (fig. 5.34)*). At the moment of transition there is a Rhamphoid Cusp but after the transition the inflection moves away from the cusp down the other branch. The cusp has been reversed and a self crossing is introduced into the RSS.

Each of these events can occur generically in a one parameter family and the conditions are:-

Event 1) Higher Inflection (2 type II inflections):
There exist two points $\gamma(t_1)$, $\gamma(t_2)$ such that

$$\begin{aligned}\kappa(t_1) &= \kappa(t_2),\\ \kappa'(t_1) &= -\kappa'(t_2) \neq 0,\\ (t_1, t_2) &\in SS_1^{-1};\end{aligned}$$

Event 2) Higher Inflection (a type I and a type II inflection):
There exist two points $\gamma(t_1)$, $\gamma(t_2)$ such that

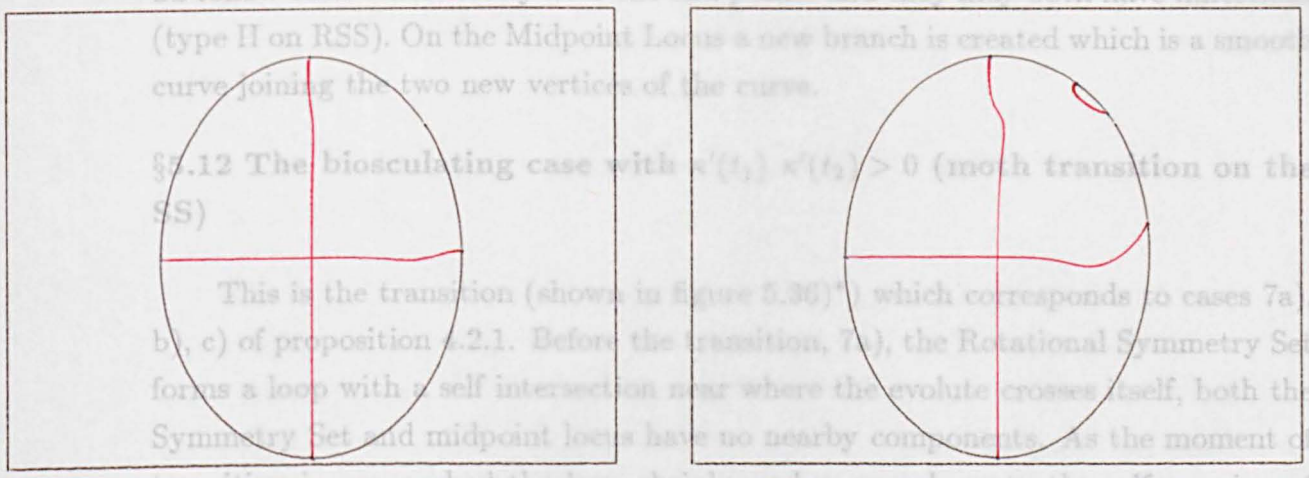
$$\begin{aligned}\kappa(t_1) &= \kappa(t_2),\\ \kappa'(t_1) &= 0,\\ \kappa'(t_2) &\neq 0,\\ (t_1, t_2) &\in SS_1^{-1};\end{aligned}$$

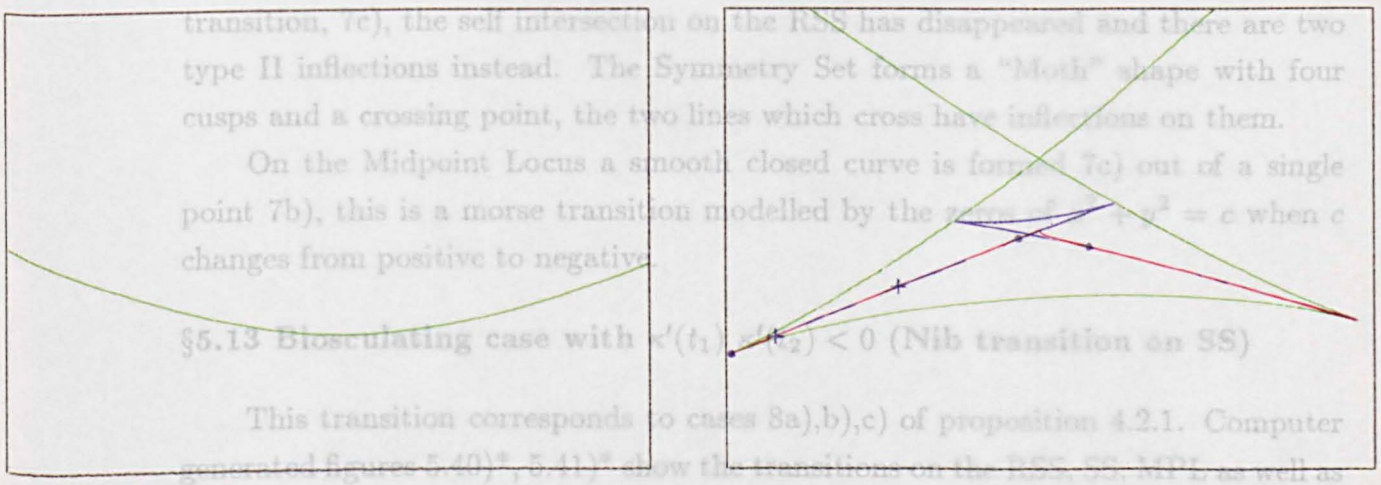
Event 3) Rhamphoid Cusp;
There exist two points $\gamma(t_1)$, $\gamma(t_2)$ such that

$$\begin{aligned}\kappa(t_1) &= \kappa(t_2),\\ \kappa'(t_1) &= \kappa'(t_2),\\ (t_1, t_2) &\in SS_1^{-1}.\end{aligned}$$

§5.11 The A_4 transition on RSS, SS and MPL

The A_4 transition on the curve occurs when two new vertices (one maximum and one minimum of curvature) are created at the same place on the curve. This produces a swallowtail transition on the evolute (each cusp of the evolute corresponds to a vertex of the curve). New branches of the RSS and the SS are created which are both contained inside the swallowtail (fig. 5.35)*). The Rotational Symmetry Set has been described in §5.3 and has one cusp, two inflections and two end-points one in each cusp of the evolute. The Symmetry Set has two cusps a self crossing and two end-points which also lie at the cusps of the evolute. The RSS and





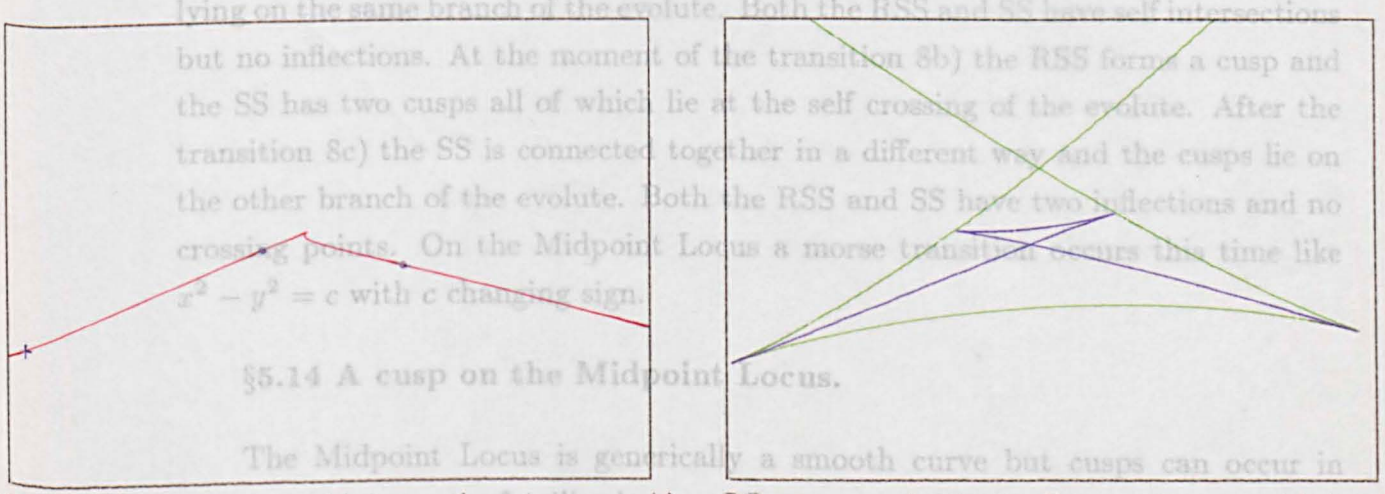


Fig 5.35) A_4 Transition

SS follow each other closely near the end points and they may both have inflections (type II on RSS). On the Midpoint Locus a new branch is created which is a smooth curve joining the two new vertices of the curve.

§5.12 The biosculating case with $\kappa'(t_1)\ \kappa'(t_2) > 0$ (moth transition on the SS)

This is the transition (shown in figure 5.36)*) which corresponds to cases 7a), b), c) of proposition 4.2.1. Before the transition, 7a), the Rotational Symmetry Set forms a loop with a self intersection near where the evolute crosses itself, both the Symmetry Set and midpoint locus have no nearby components. As the moment of transition is approached the loop shrinks and moves closer to the self crossing of the evolute. At the moment of transition, 7b), the RSS forms a cusp at the self crossing of the evolute. The Symmetry Set has and isolated point here. After the transition, 7c), the self intersection on the RSS has disappeared and there are two type II inflections instead. The Symmetry Set forms a "Moth" shape with four cusps and a crossing point, the two lines which cross have inflections on them.

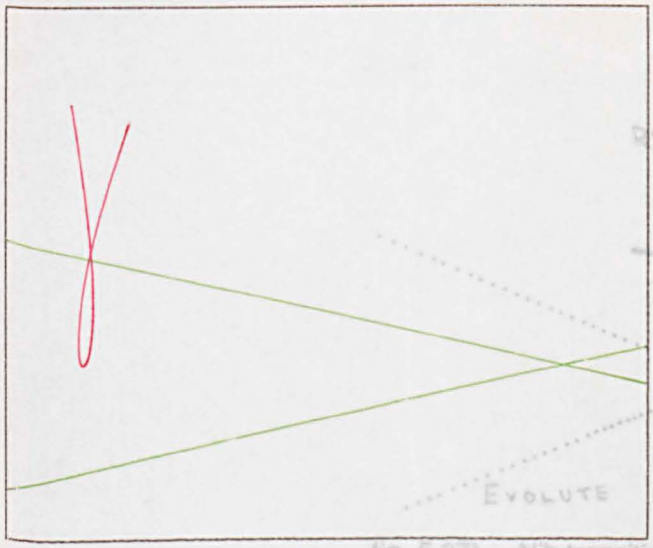
On the Midpoint Locus a smooth closed curve is formed 7c) out of a single point 7b), this is a morse transition modelled by the zeros of $x^2 + y^2 = c$ when c changes from positive to negative.

§5.13 Biosculating case with $\kappa'(t_1)\ \kappa'(t_2) < 0$ (Nib transition on SS)

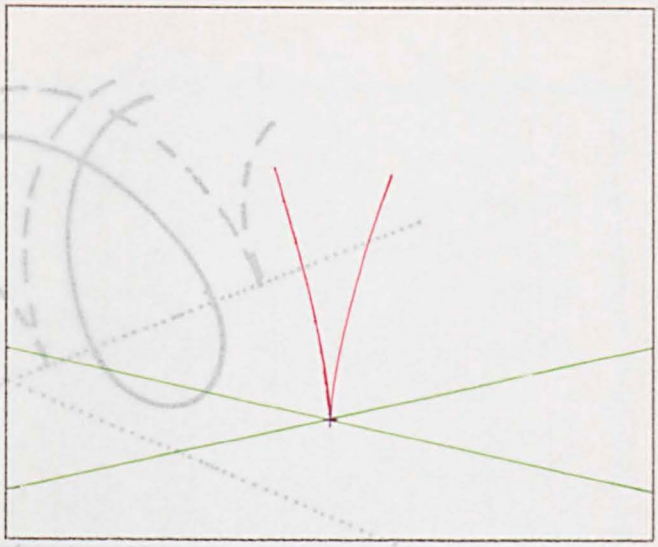
This transition corresponds to cases 8a),b),c) of proposition 4.2.1. Computer generated figures 5.40)*, 5.41)* show the transitions on the RSS, SS, MPL as well as the RSS and SS combined. Figures 5.37)–5.39) are provided to help understand the subtle geometry. Before the transition, 8a), the Symmetry Set has two cusps both lying on the same branch of the evolute. Both the RSS and SS have self intersections but no inflections. At the moment of the transition 8b) the RSS forms a cusp and the SS has two cusps all of which lie at the self crossing of the evolute. After the transition 8c) the SS is connected together in a different way and the cusps lie on the other branch of the evolute. Both the RSS and SS have two inflections and no crossing points. On the Midpoint Locus a morse transition occurs this time like $x^2 - y^2 = c$ with c changing sign.

§5.14 A cusp on the Midpoint Locus.

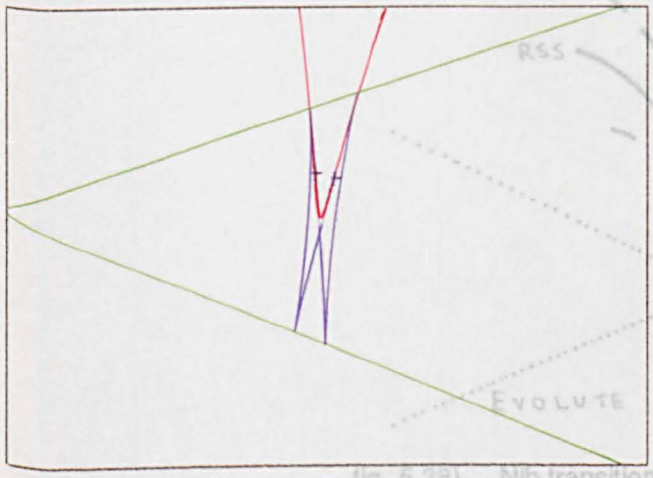
The Midpoint Locus is generically a smooth curve but cusps can occur in generic one parameter families in two different ways ([Giblin-Brassett] lemma 5.2).



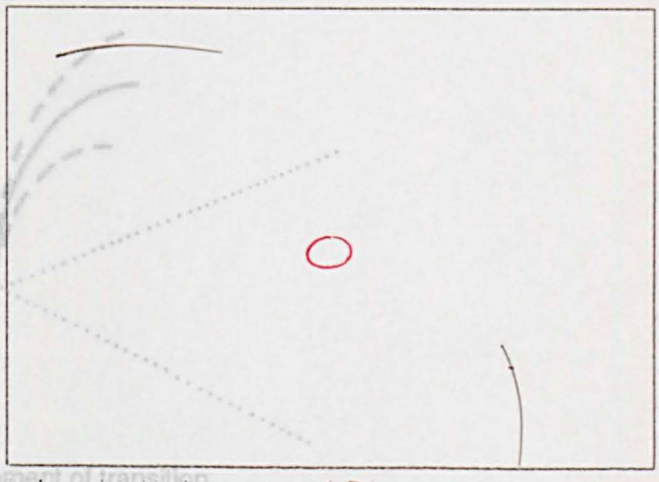
Before Transition



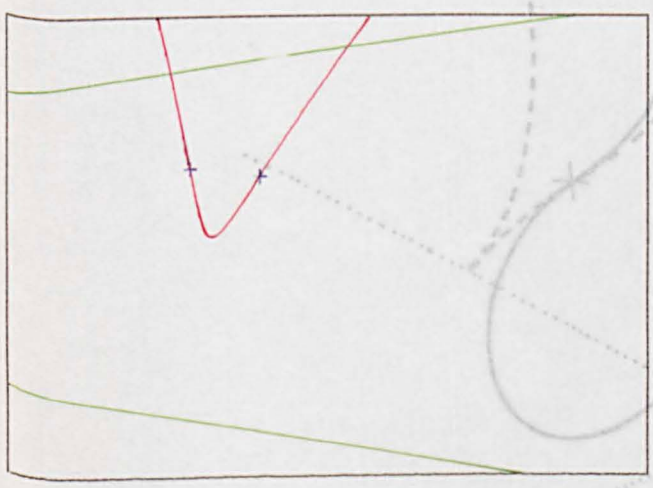
Moment of Transition



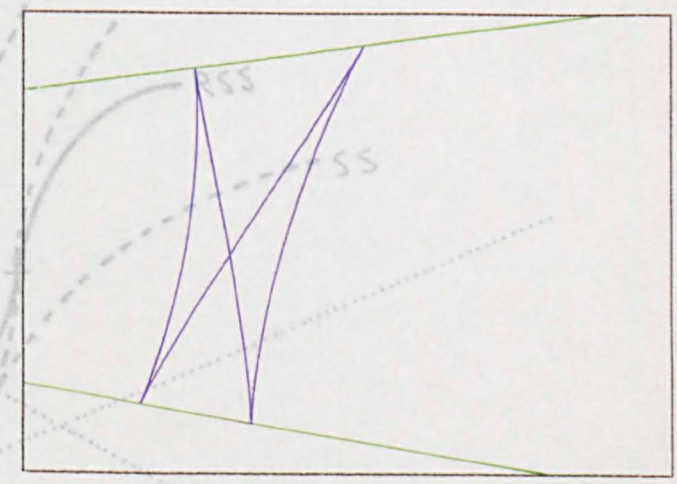
After Transition



Midpoint Locus After transition



RSS After Transition



SS After Transition (a moth)

Fig 5.36: Biosculating transition with

$$k'(t_1)\,k'(t_2) > 0$$

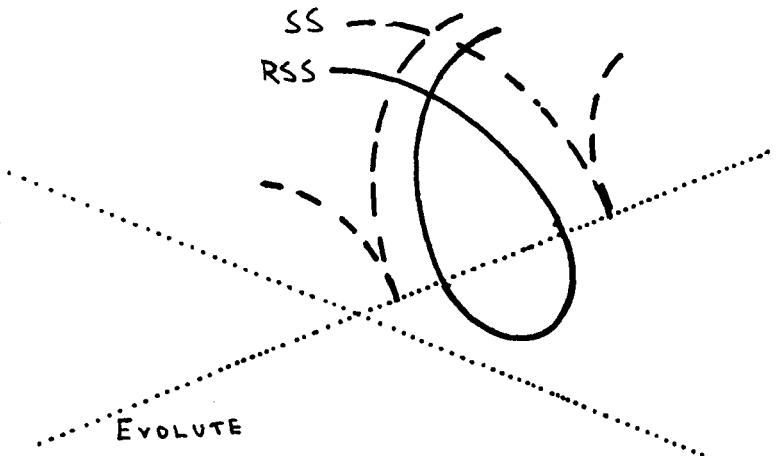


fig. 5.37) Nib transition:- before transition.

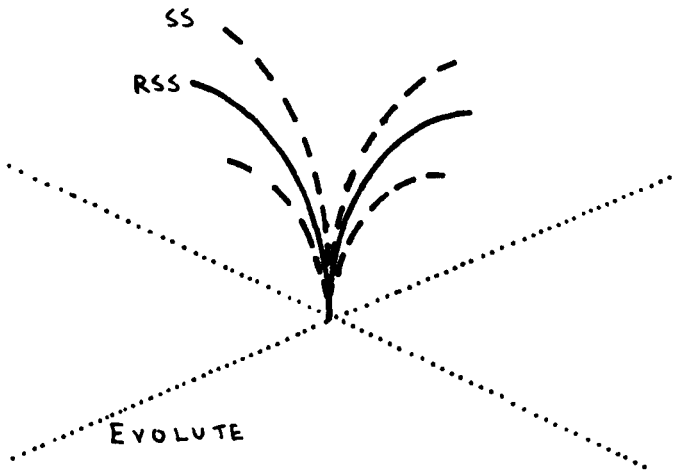


fig. 5.38) Nib transition:- moment of transition.

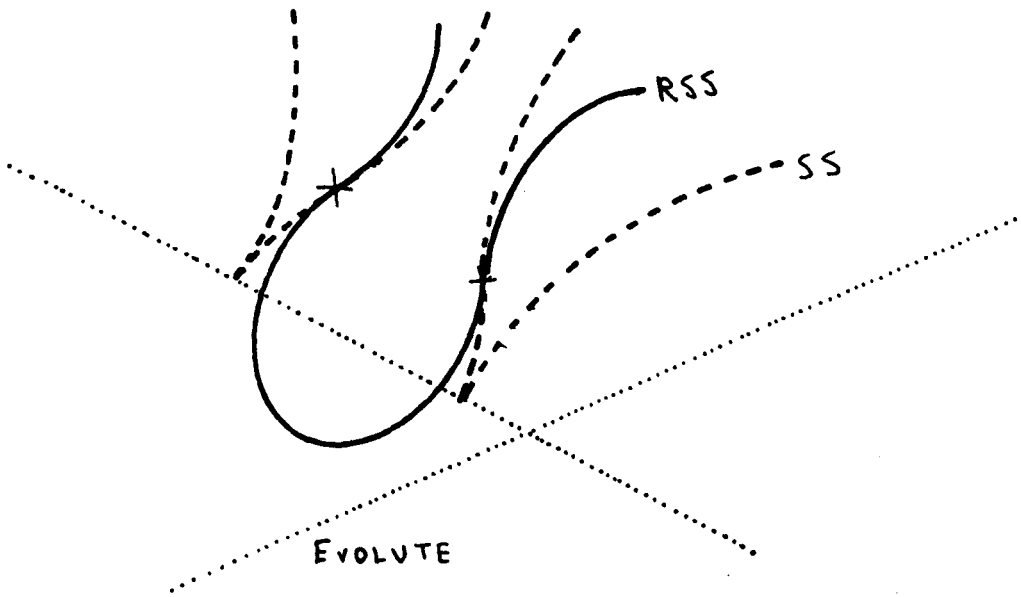


fig. 5.39) Nib transition:- after transition.

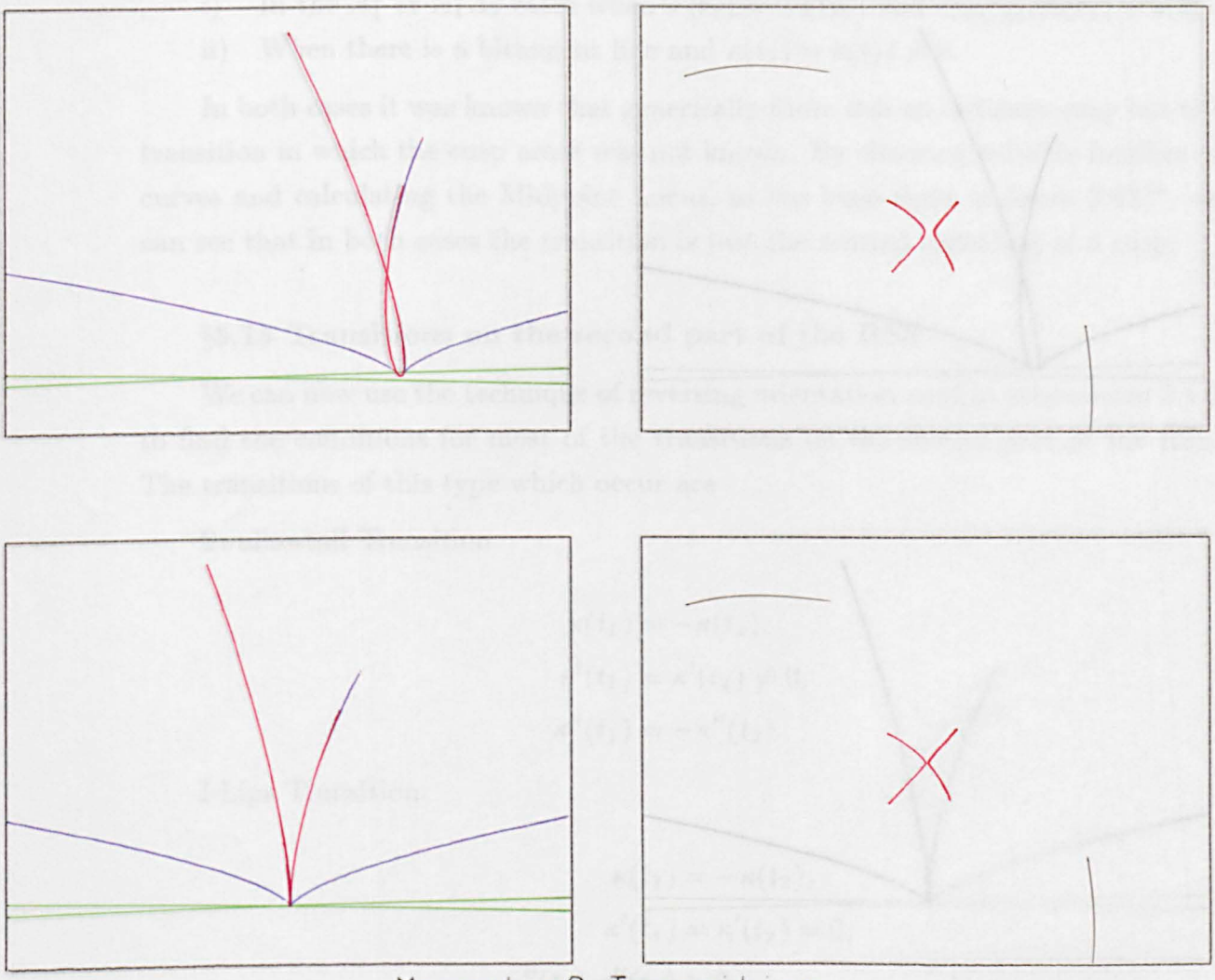


Fig 5.40) Biosculating transition with $k'(t_1)\,k'(t_2) < 0$

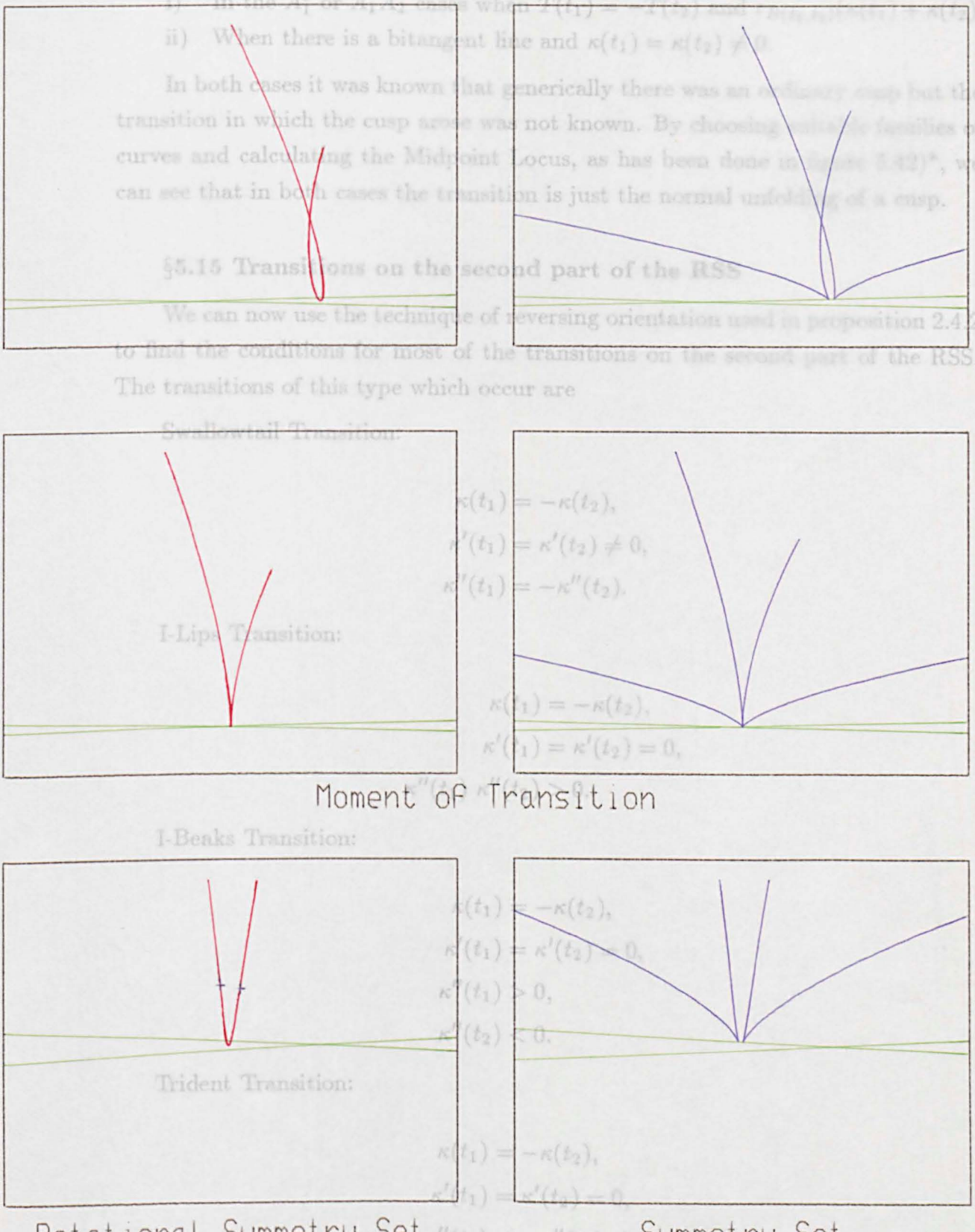


Fig 5.41) Biosculating transition with $k'(t_1)\,k'(t_2) < 0$

i) In the A_1^2 or A_1A_2 cases when $T(t_1) = -T(t_2)$ and $r_{B(t_1,t_2)}(\kappa(t_1) + \kappa(t_2)) = 2$,

ii) When there is a bitangent line and $\kappa(t_1) = \kappa(t_2) \neq 0$.

In both cases it was known that generically there was an ordinary cusp but the transition in which the cusp arose was not known. By choosing suitable families of curves and calculating the Midpoint Locus, as has been done in figure 5.42)*, we can see that in both cases the transition is just the normal unfolding of a cusp.

§5.15 Transitions on the second part of the RSS

We can now use the technique of reversing orientation used in proposition 2.4.2 to find the conditions for most of the transitions on the second part of the RSS. The transitions of this type which occur are

Swallowtail Transition:

$$\begin{aligned}\kappa(t_1) &= -\kappa(t_2),\\ \kappa'(t_1) &= \kappa'(t_2) \neq 0,\\ \kappa''(t_1) &= -\kappa''(t_2).\end{aligned}$$

I-Lips Transition:

$$\begin{aligned}\kappa(t_1) &= -\kappa(t_2),\\ \kappa'(t_1) &= \kappa'(t_2) = 0,\\ \kappa''(t_1)\ \kappa''(t_2) &> 0.\end{aligned}$$

I-Beaks Transition:

$$\begin{aligned}\kappa(t_1) &= -\kappa(t_2),\\ \kappa'(t_1) &= \kappa'(t_2) = 0,\\ \kappa''(t_1) &> 0,\\ \kappa''(t_2) &< 0.\end{aligned}$$

Trident Transition:

$$\begin{aligned}\kappa(t_1) &= -\kappa(t_2),\\ \kappa'(t_1) &= \kappa'(t_2) = 0,\\ \kappa''(t_1) &= -\kappa''(t_2) \neq 0.\end{aligned}$$

Higher Inflection (two type I):

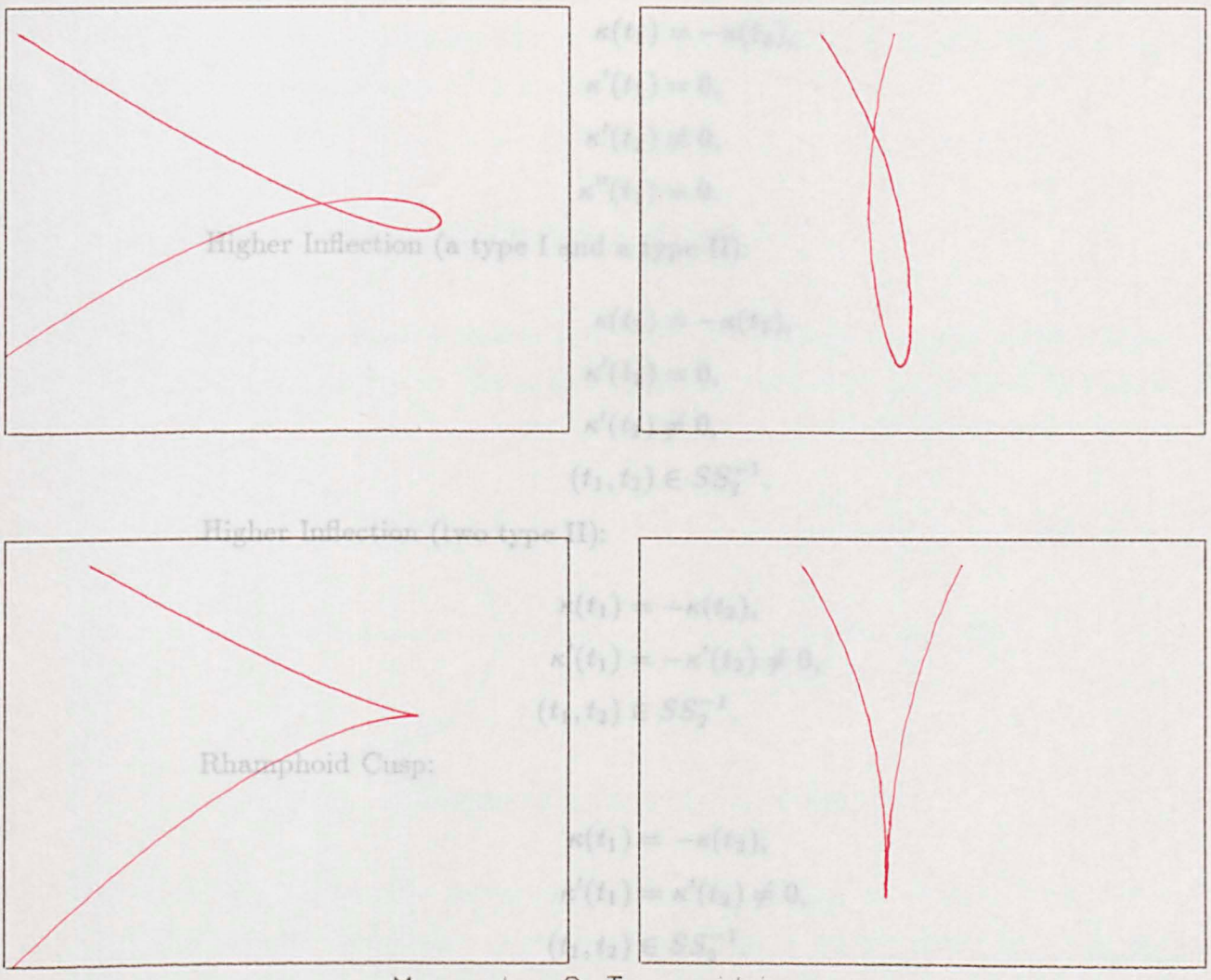


Fig 5.42) Cusp transitions on the Midpoint locus

$$\kappa(t_1) = -\kappa(t_2),$$
$$\kappa'(t_1) = 0,$$
$$\kappa'(t_2) \neq 0,$$
$$\kappa''(t_1) = 0.$$

Higher Inflection (a type I and a type II):

$$\kappa(t_1) = -\kappa(t_2),$$
$$\kappa'(t_1) = 0,$$
$$\kappa'(t_2) \neq 0,$$
$$(t_1, t_2) \in SS_2^{-1}.$$

Higher Inflection (two type II):

$$\kappa(t_1) = -\kappa(t_2),$$
$$\kappa'(t_1) = -\kappa'(t_2) \neq 0,$$
$$(t_1, t_2) \in SS_2^{-1}.$$

Rhamphoid Cusp:

$$\kappa(t_1) = -\kappa(t_2),$$
$$\kappa'(t_1) = \kappa'(t_2) \neq 0,$$
$$(t_1, t_2) \in SS_2^{-1}.$$

Triple Crossing:

$$\kappa(t_1) = \kappa(t_2) = -\kappa(t_3),$$
$$C_1(t_1, t_2) = C_2(t_1, t_3) = C_2(t_2, t_3),$$

this will involve two segments of the second part of the RSS crossing each other as well as crossing a segment of the first part of the RSS at the same point.

Unfolding of a cusp:

$$\kappa(t_1) = -\kappa(t_2),$$
$$(t_1, t_2) \in SS_2^{-1},$$

and the bitangent circle osculates at both points of contact.

The A_4 transition on γ does not create a new branch of the second part of the RSS but it can produce a number of higher inflections (first condition for higher inflections above). In addition to these there are two transitions which are unique to

the second part of the RSS. The first of these occurs when a new pair of inflections is created on γ. To deduce the shape of this new branch we will look at a slightly altered version of the κ-κ' diagram.

§5.16 The κ-κ' diagram for the RSS part 2.

To consider how we should adjust the κ-κ' diagram we again use the technique of reversing orientation. Suppose we have a normal κ-κ' diagram for two pieces of

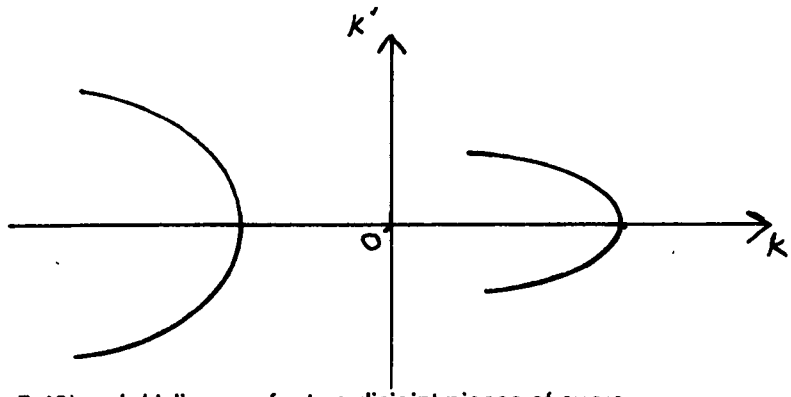


fig. 5.43) *k-k'* diagram for two disjoint pieces of curve.

curve γ_1, γ_2 where the curvatures are opposite in sign, as in figure 5.43). Now if we reverse the orientation γ_2 we change the sign of curvature but leave the sign of the derivative of curvature unchanged. This is equivalent to reflecting in the line

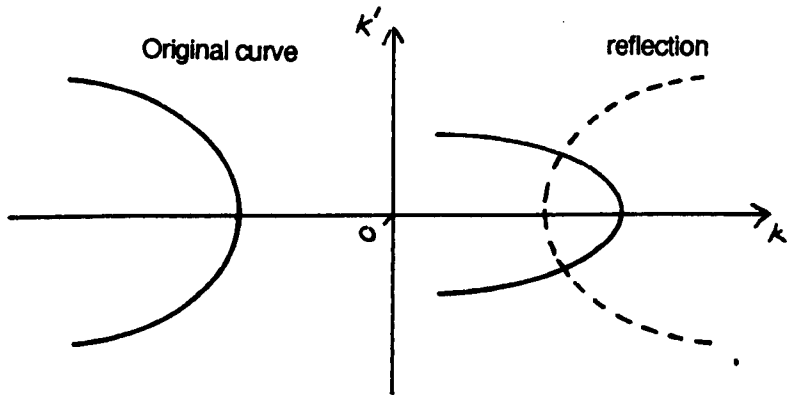


fig. 5.44) Fig. 5.43) after reflecting one arc in *k'* axis.

$\kappa = 0$ (figure 4.44). In this example we see from the κ-κ' diagram that we will have an I-lips. If we want to consider the new κ-κ' diagram for a single closed curve we reflect the whole κ-κ' diagram in the line $\kappa = 0$ (fig. 4.45). Note that since each

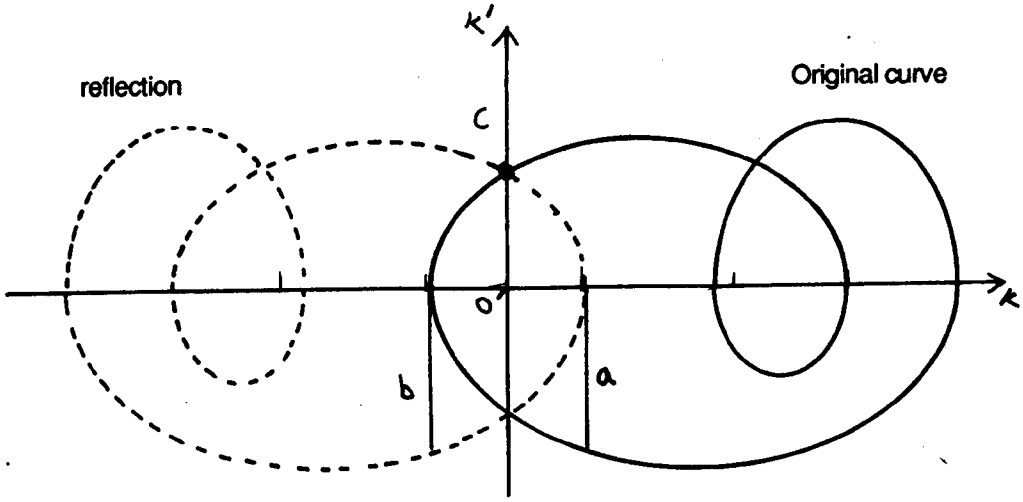


fig. 5.45) Reflecting the k-k' diagram of a closed curve in k' axis.

pair of points is represented twice, for example the pairs a and b, care should be taken to ensure that the same feature is not counted twice. Both a and b correspond to the same inflection on the RSS. The point c on the diagram does not correspond to a cusp as this point is given by a single inflection on γ and hence represents a single end point on the RSS.

§5.17 The Half Lips transition on the RSS part 2.

A pair of inflections on γ are created when the curvature of a vertex becomes zero. Figure 5.46) shows the corresponding transition on the new version of the

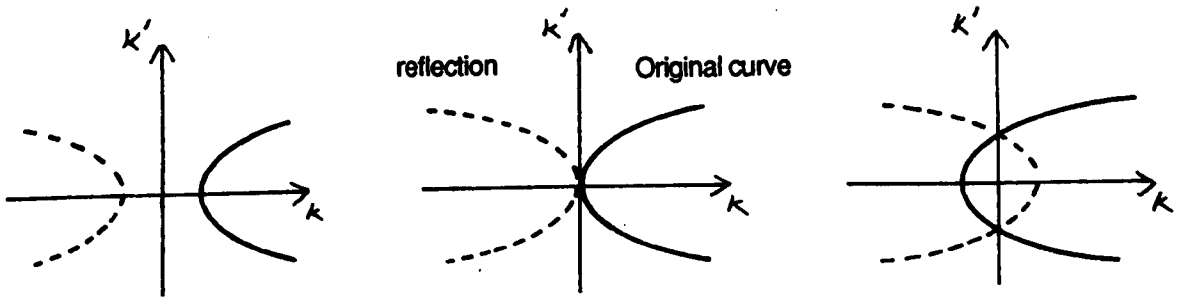


fig. 5.46) Half-Lips transition:- k-k' diagram.

κ-κ' diagram. We see that before the transition there are no pairs (t_1, t_2) with $\kappa(t_1) = -\kappa(t_2)$ so the second part of the RSS is the empty set. At the moment of transition there is one value for t for which $\kappa(t) = -\kappa(t) = 0$ and the RSS part

2 consists of an isolated point which lies at the higher inflection of the curve (fig 5.49a)*).

By following pairs of points on the new κ-κ' diagram we can discover the shape of the new part of the RSS (fig 5.47)). Starting at the end point (1) we pass through two inflection (3), (6) and arrive at an other end point (8). At the two end points

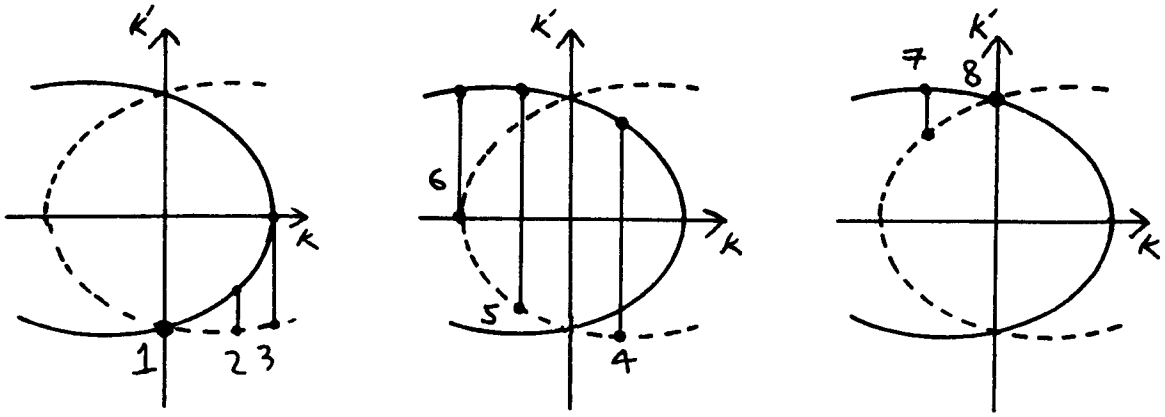


fig. 5.47) Half-Lips transition:- tracing out pairs of points on *k-k'* diagram.

the RSS is tangent to the curve and together γ and the RSS form a half lips shape as shown in the third picture in figure 5.49a)*.

§5.18 Half Cusp Transition.

If we look at the new κ-κ' diagram for a single curve where $\kappa(t) = \kappa''(t) = 0$ and the diagrams for slightly perturbed curves we see the transition illustrated in

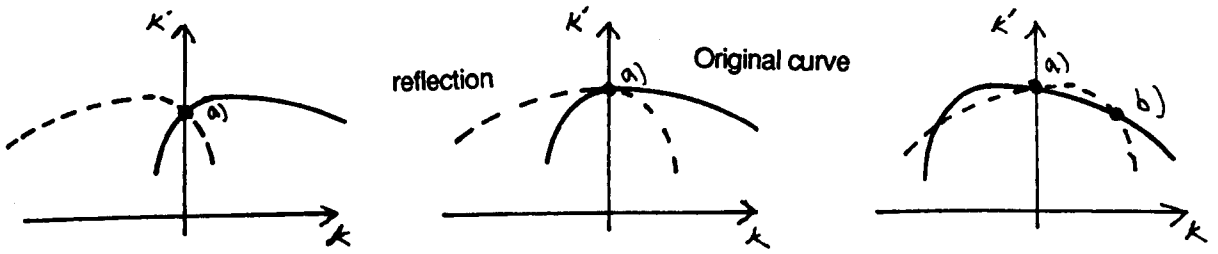


fig. 5.48) Half-Cusp transition:- *k-k'* diagram.

figure 5.48). We observe in the third picture there is a crossing point, *b*), but there is none in the first picture. At the moment of transition the crossing point is absorbed in the end point *a*). On the Rotational Symmetry Set we see that in the transition a cusp grows out of the inflection on the curve (fig. 5.49b)*). At all times in the transition the curve and the end of the RSS are tangent but the direction of the tangent to the RSS is reversed by 180 degrees during the transition.

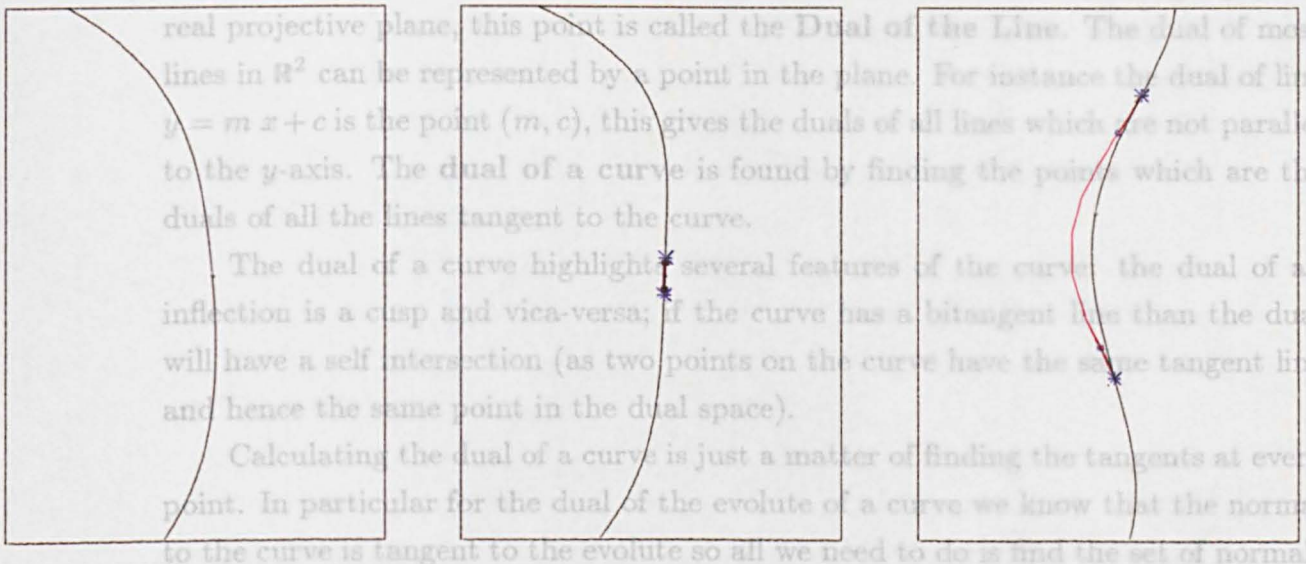


Fig 5.49a) Half Lips transition on RSS part 2

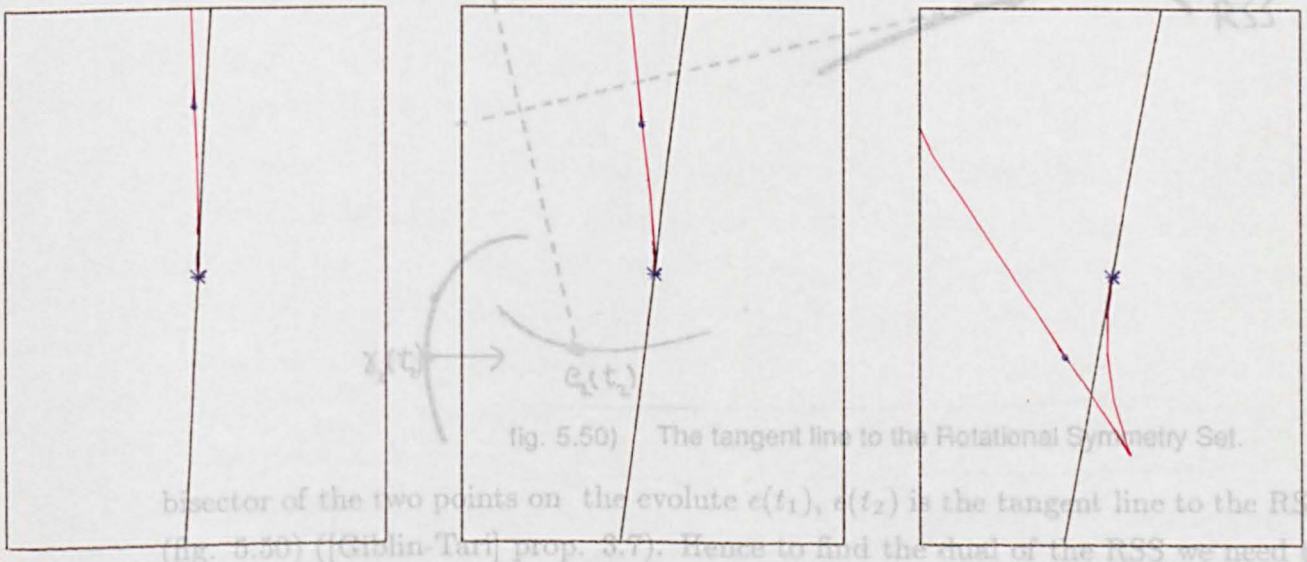


Fig 5.49b) Half cusp transition on RSS part 2

§5.19 Duals of transitions

Any line in the real projective plane can be represented by a point lying in the real projective plane, this point is called the **Dual of the Line**. The dual of most lines in $\mathbf{R}^2$ can be represented by a point in the plane. For instance the dual of line $y = m\,x + c$ is the point (m, c), this gives the duals of all lines which are not parallel to the y-axis. The **dual of a curve** is found by finding the points which are the duals of all the lines tangent to the curve.

The dual of a curve highlights several features of the curve: the dual of an inflection is a cusp and vica-versa; if the curve has a bitangent line than the dual will have a self intersection (as two points on the curve have the same tangent line and hence the same point in the dual space).

Calculating the dual of a curve is just a matter of finding the tangents at every point. In particular for the dual of the evolute of a curve we know that the normal to the curve is tangent to the evolute so all we need to do is find the set of normals to the curve. To find the dual of the Rotational Symmetry Set we note that for each pair of points $\gamma(t_1)$, $\gamma(t_2)$ which gives a point on the RSS the perpendicular

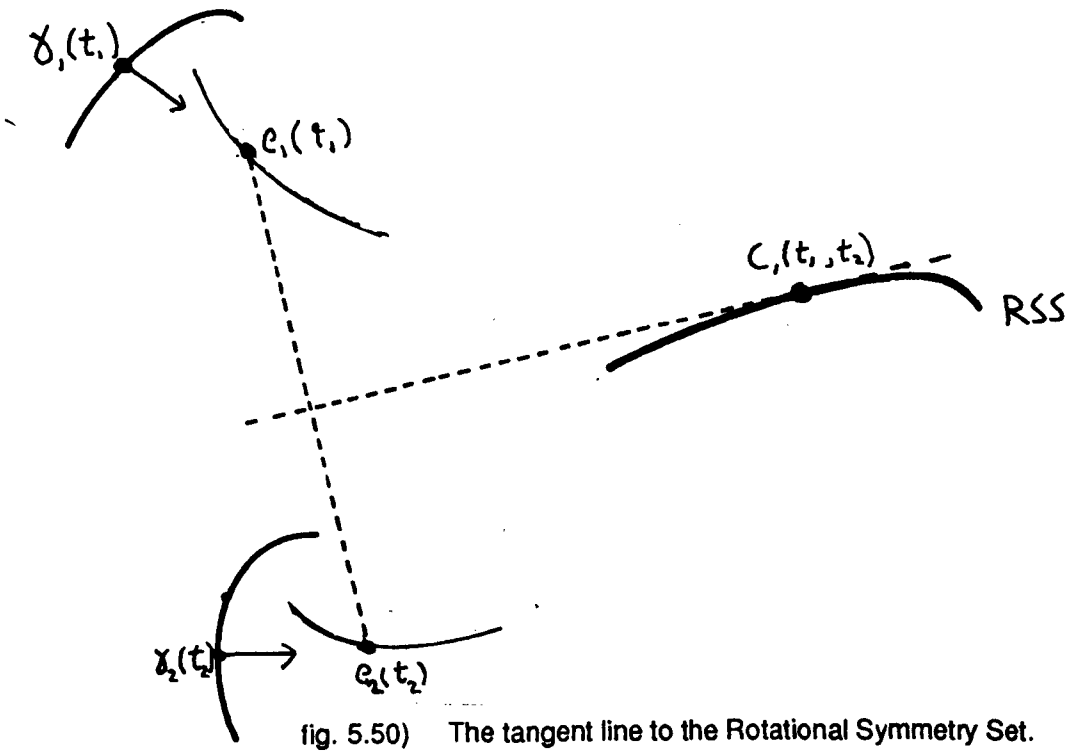


fig. 5.50) The tangent line to the Rotational Symmetry Set.

bisector of the two points on the evolute $e(t_1)$, $e(t_2)$ is the tangent line to the RSS (fig. 5.50) ([Giblin-Tari] prop. 3.7). Hence to find the dual of the RSS we need to find the duals of the lines

$$\{\ \tfrac{1}{2}(e(t_1) + e(t_2)) + \lambda\ i(e(t_1) - e(t_2))\ |\ \lambda \in \mathbf{R}\ \}.$$

If (t_1, t_2) is in SS_1^{-1} then both points $C_1(t_1, t_2)$ and $\frac{1}{2}(\gamma(t_1) + \gamma(t_2))$ will lie on the local axis of reflectional symmetry which is the tangent line to the Symmetry Set

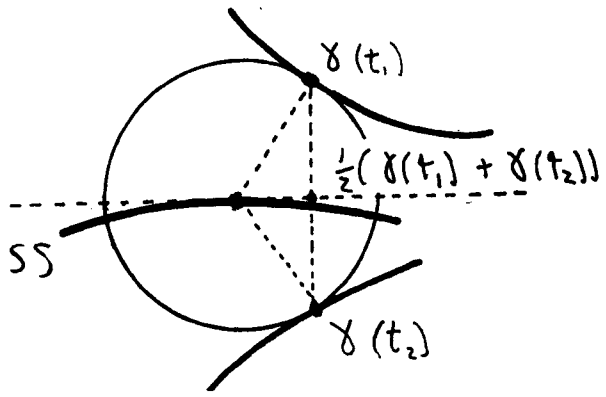


fig. 5.51) The tangent line to the Symmetry Set.

([Giblin-Brassett] §4.1). By finding the duals of the lines

$$\{\ C_1(t_1, t_2) + \lambda(\tfrac{1}{2}(\gamma(t_1) + \gamma(t_2)) - C_1(t_1, t_2))\ |\ \lambda \in \mathbf{R}\ \}$$

we will find the dual of the symmetry set. If $C_1(t_1, t_2) = \frac{1}{2}(\gamma(t_1) + \gamma(t_2))$ the above set does not define a line. However we know that the two pieces of curve will have parallel tangent lines which is also parallel to the tangent to SS so we can calculate the dual here. At a vertex of the curve the RSS and SS are both tangent to the normal to the curve and at an inflection the second part of the RSS is tangent to the curve. We can use these facts to find the duals of these sets as t_1 and t_2 both approach the same value.

By examining some of the computer pictures here a striking similarity between the Rotational Symmetry Set and the dual of the Symmetry Set (or equivalently the SS and the dual of the RSS) was first observed. This is particularly noticeable for transitions. This has led to an indepth investigation [Giblin-Tari], [Tari]. An example of this similarity can be seen in the dual of the I-Beak transition on the RSS (fig 5.26)*) which resembles the Nib transition on the SS (fig. 5.41)*). Also the Moth transition on the SS (fig. 5.36)*) is similar to the dual of the I-lips transition on the RSS (fig. 5.52)*). In particular the Moth and Nib both have self intersections so the dual have bitangent lines which can be seen. Also note that the moth has four cusps so its dual has four inflections which is not the standard Lips which occurs as a discriminant.

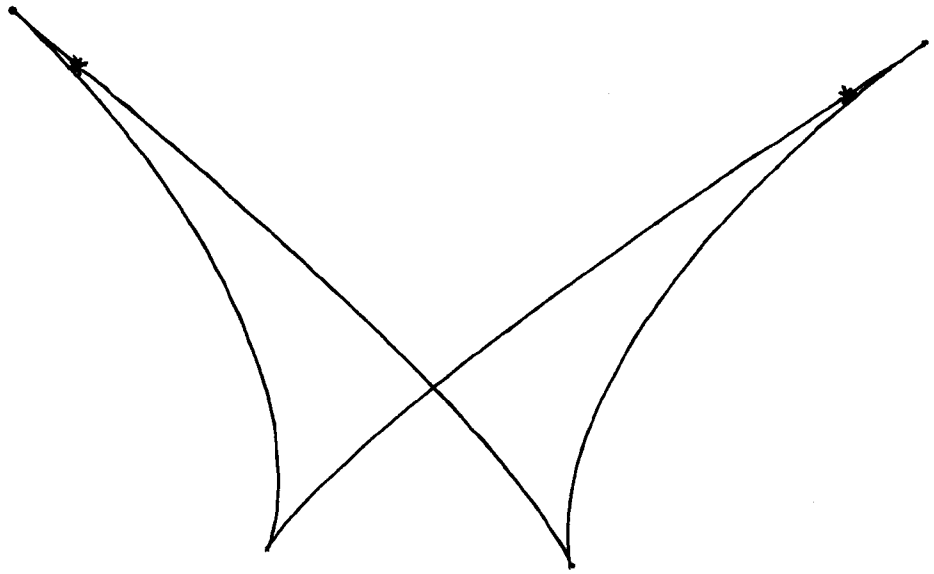


fig. 5.52)* Dual of an I-lips

If we consider the dual of the SS combined with the evolute we find it resembles part II of the RSS combined with part of the curve. In particular taking the dual of the SS and evolute in the A_4 transition gives the picture shown in figure 5.53)*. This is like the semi-lips transition on part II of the RSS.

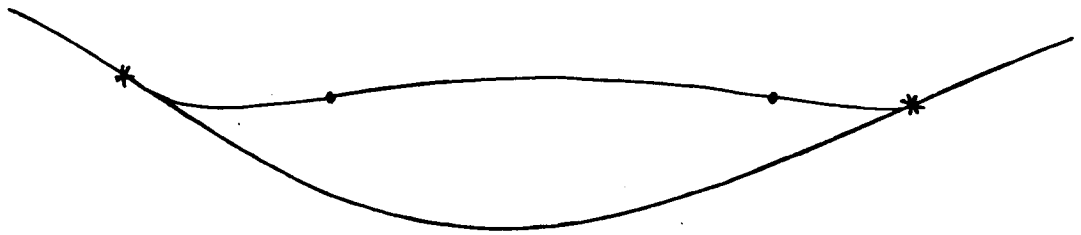


fig. 5.53)* Dual of Symmetry Set and evolute after A_4 transition.

§5.20 Piecewise Circular Curves

So far we have looked at generic curves which have continuous curvature which is not locally constant, here called an*ordinary* curves. However it is possible to consider the Symmetry Set of a curve which is made out of pieces of circle joined end to end to produce a C^1 differentiable curve. Here we review some results of [Giblin-Banchoff] about the SS of such curves and describe how the SS can be calculated.

Consider such a curve, P, made out of four circular arcs a_1, a_2, a_3, a_4 with radii r_1, r_2, r_3, and r_4 respectively. The curvature of this curve will be a step function. A circle tangent to the curve at the point between two arcs is said to be *osculating* if the value of its curvature lies between the curvatures of the two arcs. This implies

that the circle will be tangent to both arcs at their end point and also cross the curve P. This is what we would expect to happen when a circle has A_3 contact with an *ordinary* curve. The evolute of an *ordinary* curve consists of the centres of osculating circles. If we take the evolute to be the centres of all the *osculating* circles here we find that the evolute will be a polygon joining the centres of the

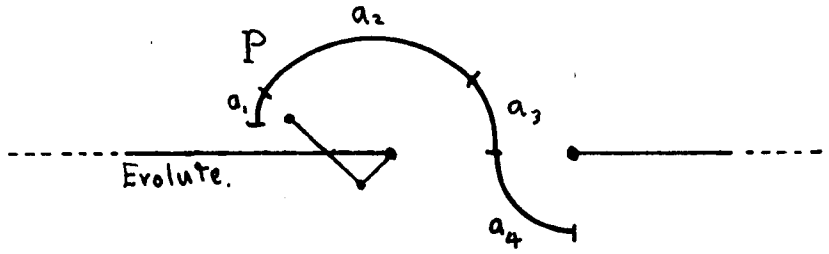


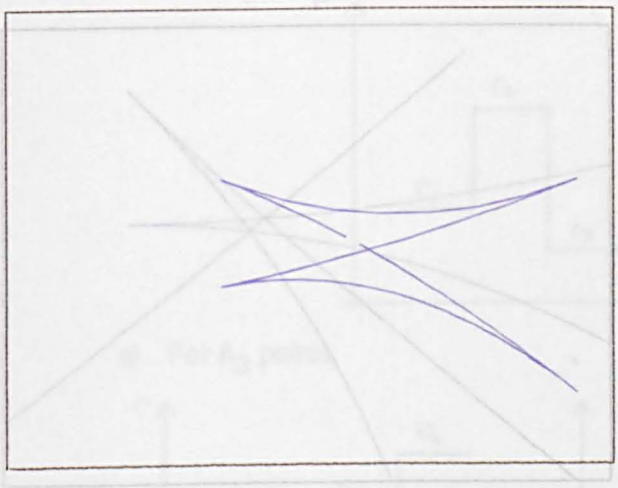
fig. 5.54) A piecewise circular curve and its evolute.

circular arcs (fig.5.54). If, as between the arcs a_3 and a_4 in this figure ,there is an inflection then the evolute will join the two points via the point at infinity. This point at infinity corresponds to the osculating line which is crosses P and is tangent to P between arcs a_3 and a_4.

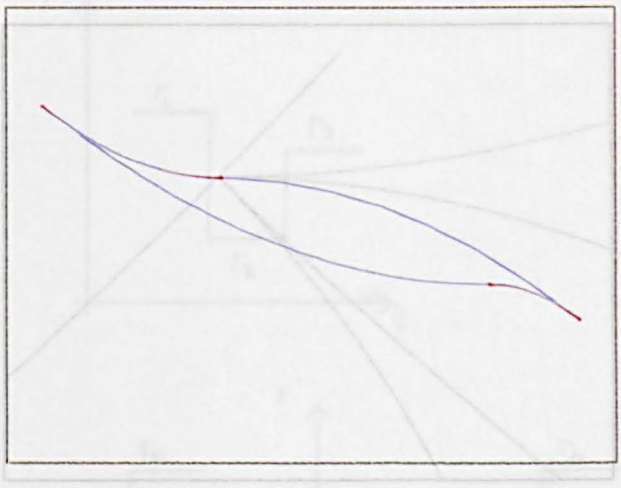
It can be shown [Giblin-Banchoff] that the Symmetry Set of such a curve is made up of a number of conic sections (ellipses, hyperbolas, parabolas and straight lines) which are joined end to end to form *inflections*, *"cusps"* and *"rhamphoid cusps"*. *Cusps* have one branch on each side of the tangent line and *rhamphoid cusps* have both branches on the same side. Here *rhamphoid cusps* occur generically and can be regarded as a combination of an *inflection* and a *cusp* (they are non generic for an *ordinary* curve).

Many of the transitions which occur on the SS of an *ordinary* curve have their equivalent on the Symmetry Set of a Piecewise Circular Curve. Some of these transitions occur in two different ways,one involving cusps and the other with rhamphoid cusps. For example in the Moth transition, fig 5.55)*, we have two different forms and the duals are also different with the bitangent line in different places. Also illustrated, figure 5.56)*, are two forms of the $A_1A_1A_2$ transition where a tri-tangent circle osculates at one of its points of contact.

We can think of a vertex as local extremum of curvature, hence the curve P above is said to have a "vertex" along the arc A_2 if $r_2 > r_1$ and $r_2 > r_3$ or if $r_2 < r_1$ and $r_2 < r_3$ (fig. 5.57). The evolute of the Piecwise Circular Curve near a vertex will form a corner which is also an ending point of the Symmetry Set. For the purpose of calculating the SS we would start following the h^{-1} contour from the point (j_1, j_2) where j_1 and j_2 are the end points of the arc A_2. Figure 5.58)* shows

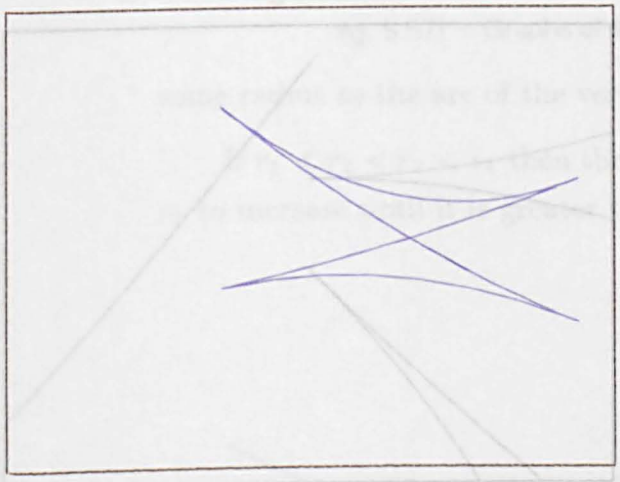


Symmetry Set

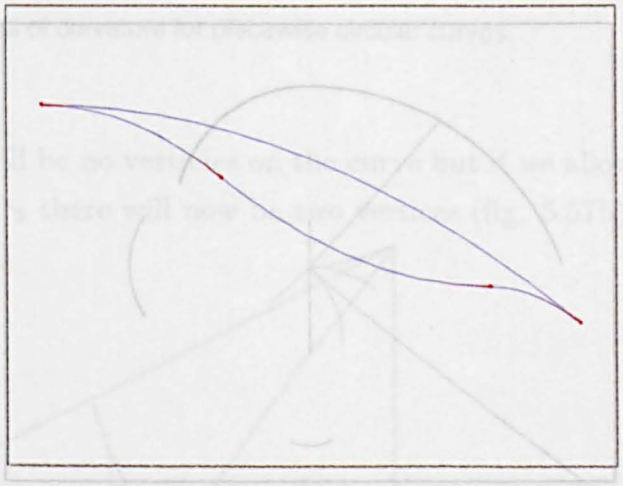


Dual

Example 1



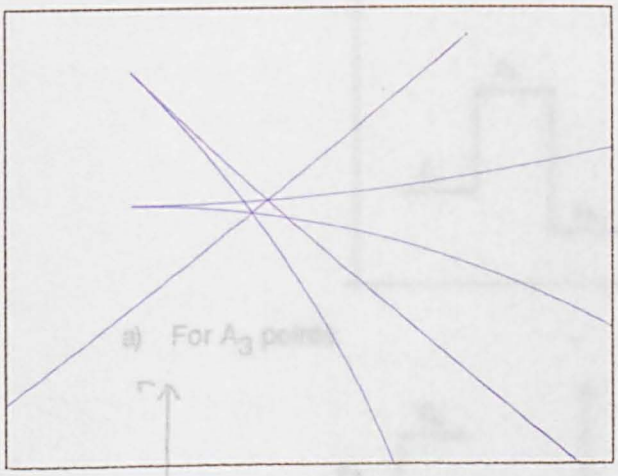
Symmetry Set



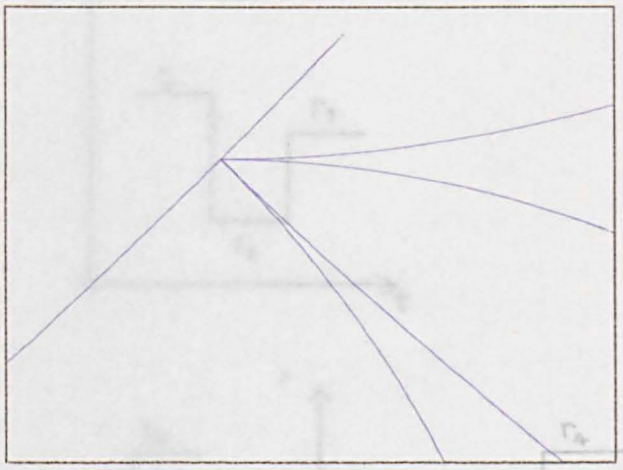
Dual

Example 2

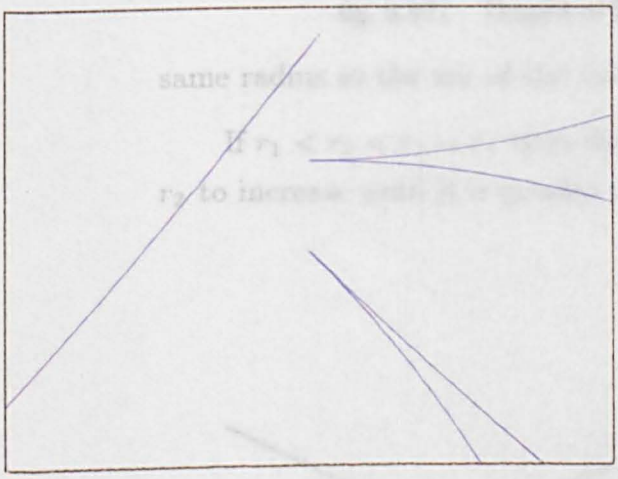
Fig 5.55: Two examples of the Moth transition for piecewise circular curves



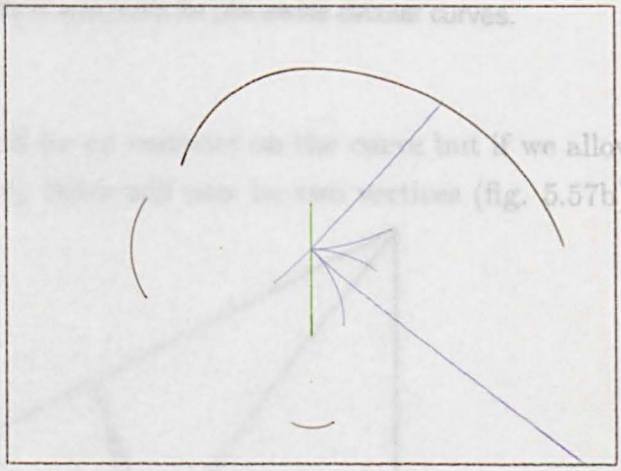
Before transition



Moment of transition



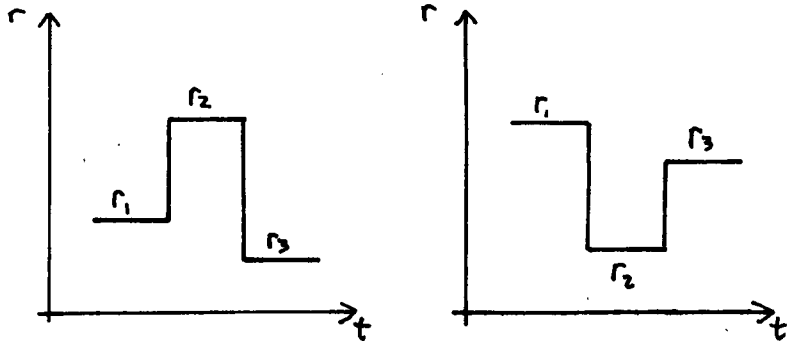
After transition



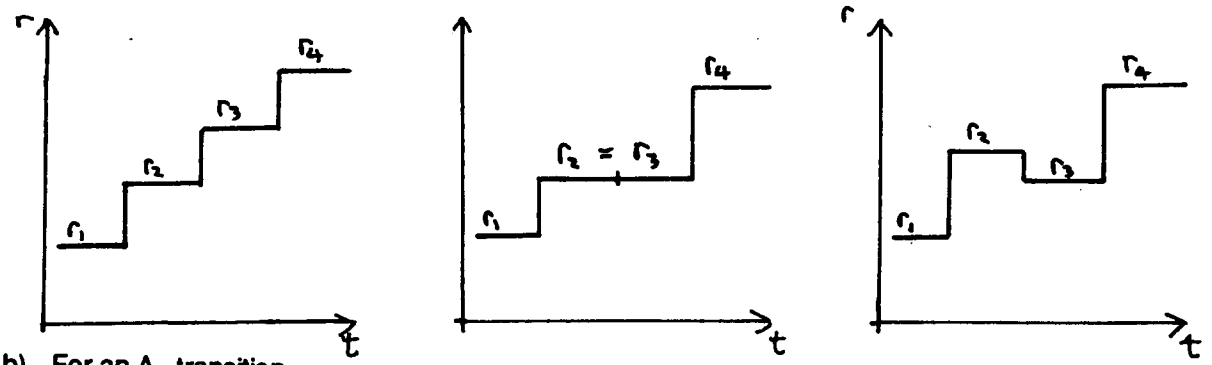
The curves used

Fig 5.56: $A_1A_1A_2$ transition on a piecewise circular curve

two forms of the A_1A_3 transition (a circle is said to have A_3 contact if it has the



a) For A_3 points.



b) For an A_4 transition.

fig. 5.57) Graphs of the radii of curvature for piecewise circular curves.

same radius as the arc of the vertex).

If $r_1 < r_2 < r_3 < r_4$ then there will be no verticies on the curve but if we allow r_2 to increase until it is greater than r_3 there will now be two vertices (fig. 5.57b).

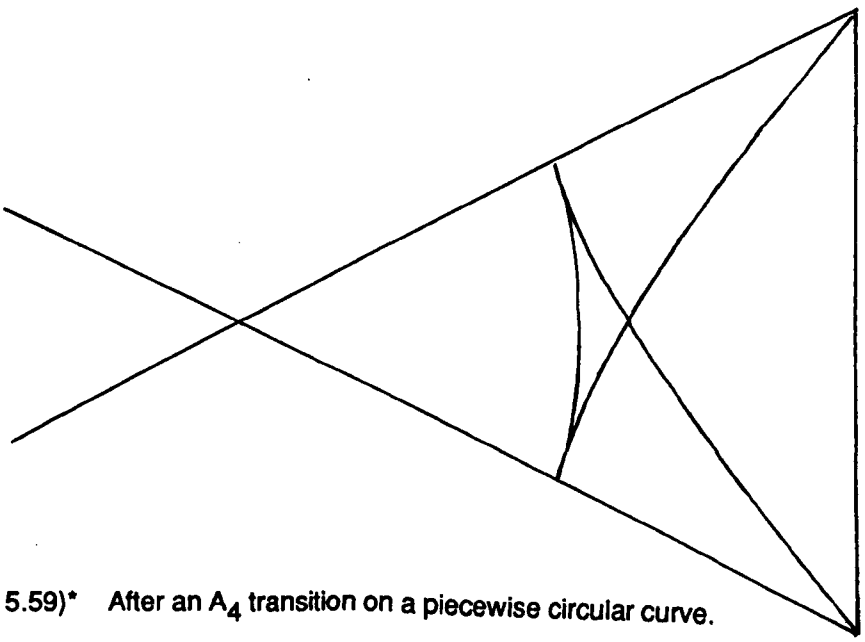
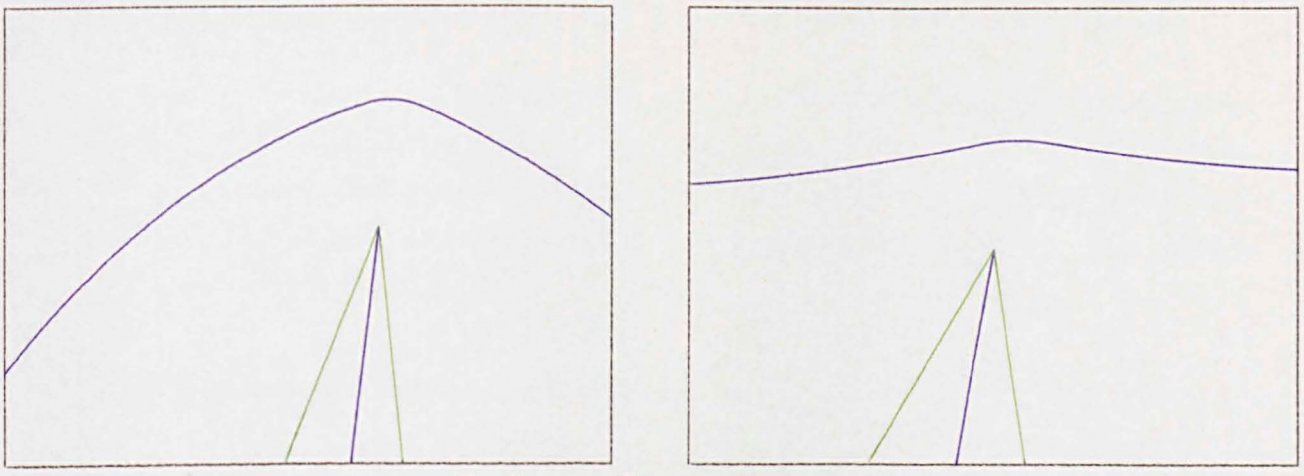


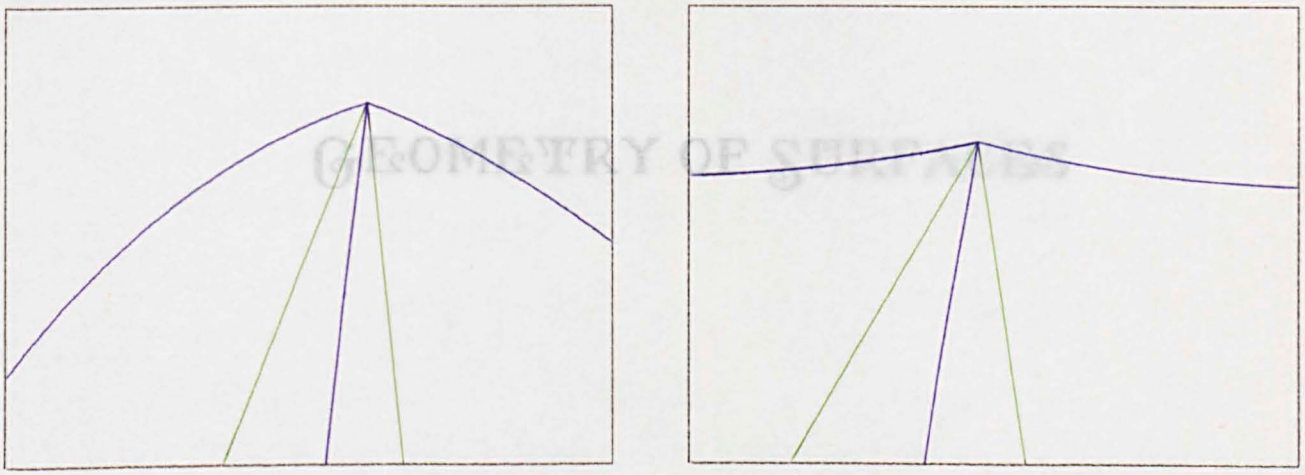
fig. 5.59)* After an A_4 transition on a piecewise circular curve.

This transition can be regarded as an A_4 transition and the SS of such a transition

is illustrated in fig 5.59)*.

A few changes to the program are necessary for it to cope with this new class of curves. The curvature is a step function and its derivative is taken to be an impulse function, which is zero in the middle of an arc and assumes a fixed positive or negative value near the ends of the arc. The sign depends on the curvatures of the two arcs. This type of function is necessary to enable the test for osculating circles to work. Also needed is a different representation of a vertex, instead of a single point for *ordinary* curves they must be considered as a pair of points one at each end of the arc. As only a few examples of vertices were required the coordinates were hand entered in each example. This saved some time in writing a routine to find them automatically.





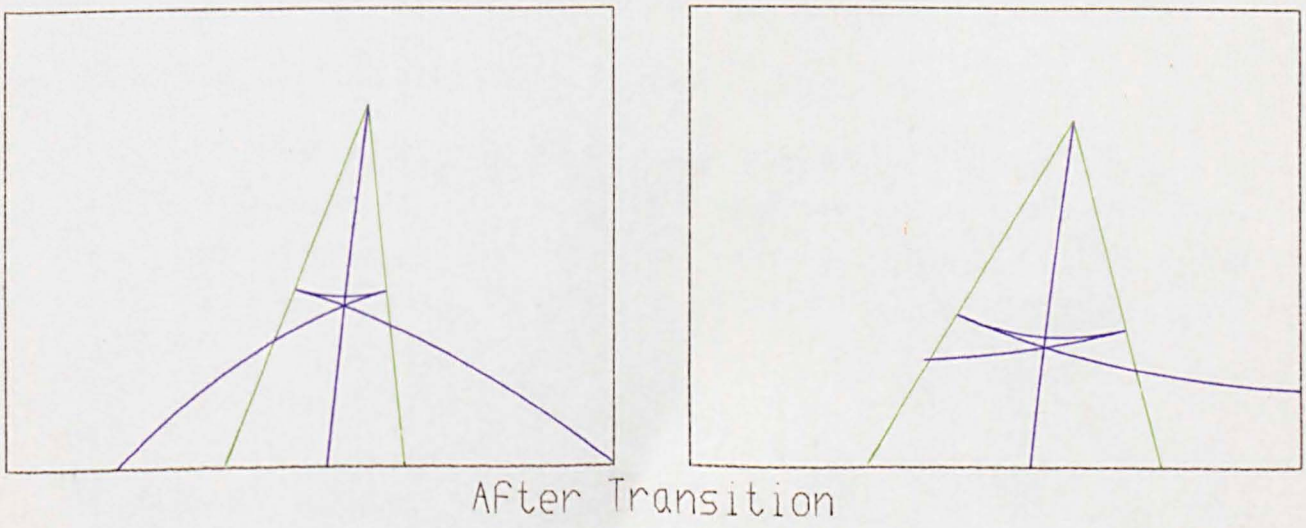


Fig 5.58: Two forms of the A_1A_3 transition for piecwise circular curves

PART 2

GEOMETRY OF SURFACES

Chapter 6: Introduction to the geometry of surfaces

In this chapter we review some standard results about the geometry of smooth surfaces in $\mathbf{R}^3$ and their associated focal surfaces. We also look at the singularities of focal surfaces. For a more intuitive discussion of some of these topics see [Koenderink]. The mathematics is discussed in more detail in many places for example [O'Neill] and [do Carmo].

§6.1 Surfaces and derivatives of maps

We start by defining a surface.

Definition 6.1

Let U be an open set in $\mathbf{R}^2$ and let s be a smooth embedding of U in $\mathbf{R}^3$. The image $S = s(U)$ is call a **surface patch** and the map s is called a **parametrization of the surface**. The parametrization is not unique to the surface.

The parametrization will normally be given in **Monge form** i.e.

$$s : (x_1, x_2) \to (x_1, x_2, ax_1^2 + bx_1x_2 + cx_2^2 + \text{higher order terms}).$$

We can always move a surface by translation and rotation such that any desired point lies at the origin and it tangent plane is the $z = 0$ plane. Such a surface can be parametrized in Monge form. Furthermore we can then rotate around the z-axis so that the quadratic term becomes $ax_1^2 + cx_2^2$.

Definition 6.2 The derivative of a map

Let U, V be open sets in $\mathbf{R}^m$, $\mathbf{R}^n$ respectively and let s be a map from U to V. The **Derivative of s at a point $x \in U$** is a linear map $ds_x : \mathbf{R}^m \to \mathbf{R}^n$ which sends tangent vectors to U at x to tangent vectors to V at $s(x)$. If $a : \mathbf{R} \to U$ is a curve with $x = a(0)$ then $u = \frac{da}{dt}|_{t=0}$ is a tangent vector to U. The map ds_x is defined by

$$ds_x<u> = \left.\frac{d(s \circ a)}{dt}\right|_{t=0}.$$

Angle brackets are used to indicate that the argument is a tangent vector and that the map is linear. In particular for a surface in $\mathbf{R}^3$ we have $m = 2$, $n = 3$ and the derivative of the map is as illustrated in figure 6.1.

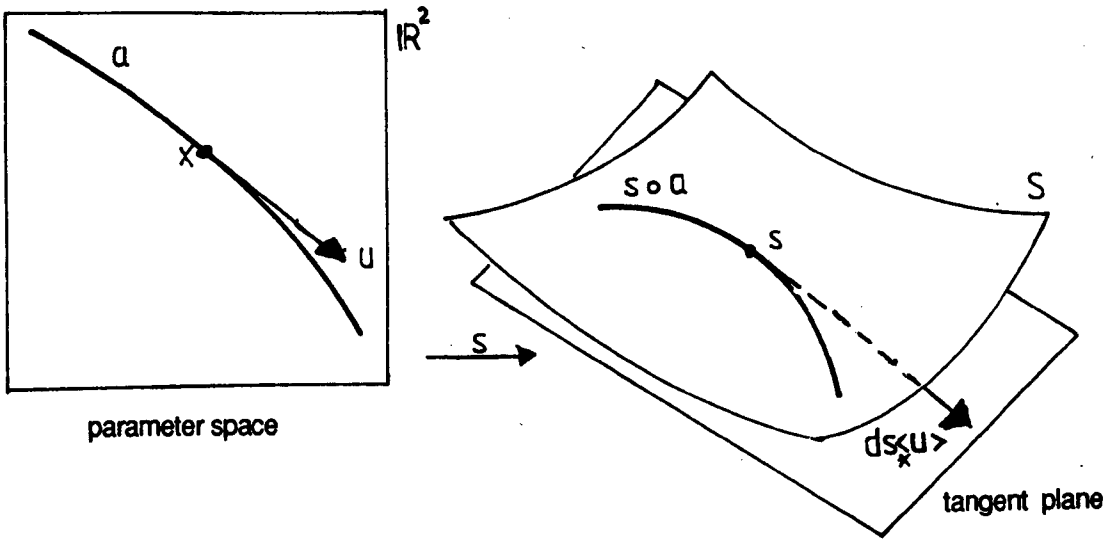


fig. 6.1) The derivative of a map s: $\mathbb{R}^2 \to \mathbb{R}^3$

Let $[A, B]$ denote the set of linear maps from A to B. The map ds_x is an element of $[\mathbf{R}^m, \mathbf{R}^n]$ so we can construct a map $ds : U \to [\mathbf{R}^m, \mathbf{R}^n]$ by defining $ds(x) = ds_x$. We can differentiate this map at a point x to give a linear map

$$d^2 s_x \in [\mathbf{R}^m, [\mathbf{R}^m, \mathbf{R}^n]] \cong [\mathbf{R}^m \times \mathbf{R}^m, \mathbf{R}^n].$$

This map is called the **second derivative of** s **at** x and is symmetric and bi-linear map. Similarly we can construct the second derivative of s

$$d^2 s : U \to [\mathbf{R}^m \times \mathbf{R}^m, \mathbf{R}^n],$$

and all the higher derivatives $d^3 s$, $d^4 s$ etc.

These derivatives can be expressed in terms of their various partial derivatives. If

$$s : (x_1, x_2) \to (X(x_1, x_2), Y(x_1, x_2), Z(x_1, x_2))$$

is a parametrization of a surface and if $u = (a, b)$, $v = (c, d)$ are tangent vectors to $\mathbf{R}^2$ then the first and second derivatives at a point $x = (x_1, x_2)$ are

$$ds_x<u> = \left(a \left.\frac{\partial X}{\partial x_1}\right|_x + b \left.\frac{\partial X}{\partial x_2}\right|_x, \ldots \right)$$

$$d^2 s_x<u, v> = \left(ac \left.\frac{\partial^2 X}{\partial x_1^2}\right|_x + (ad + bc) \left.\frac{\partial^2 X}{\partial x_1 \partial x_2}\right|_x + bd \left.\frac{\partial^2 X}{\partial x_2^2}\right|_x, \ldots \right).$$

§6.2 The first and second fundamental forms, the Gauss map and principal directions.

Let s be a parametrization of a smooth surface S in $\mathbf{R}^3$. The first and second fundamental forms of s are bilinear maps which contain much of the information about the first and second derivatives of s. The second fundamental form at a point x describes the way the normal to the surface at $s(x)$ changes as we move in any direction through x. There are generally two directions in which this change is particularly simple. These directions are called the principal directions.

Definition 6.3

The **first fundamental form** of s at a point $x \in \mathbf{R}^2$ is a symmetric bilinear map

$$I_x : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$$

given by

$$I_x<u, v> = ds_x<u> \cdot ds_x<v>.$$

If $I_x<pp> = I_x<qq> = 1$ and $I_x<pq> = 0$ then $\underline{P} = ds_x<p>$, $\underline{Q} = ds_x<q>$ are unit length orthogonal vectors.

At any point, $s(x)$, of the surface there is a normal line which passes through $s(x)$ and is normal to the tangent plane at $s(x)$. There are two points which are a unit length along this line $s(x) + \underline{n}$, $s(x) - \underline{n}$. The vector $\underline{n}$ is said to be a **unit normal**. For a surface patch we can find a continuous field of unit normals $\underline{n} : \mathbf{R}^2 \to \mathbf{R}^3$. We can identify each unit normal with a point on the unit sphere. This gives a map $g : \mathbf{R}^2 \to S^2$ which is called the **Gauss Map**.

Like any other map we can differentiate $\underline{n}$ to give a linear map $d\underline{n}_x : \mathbf{R}^2 \to \mathbf{R}^3$.

Definition 6.4

The second fundamental form of s is a symmetric bilinear map

$$II_x : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$$

$$II_x<u, v> = -d\underline{n}_x<u> \cdot ds_x<v>.$$

As $\underline{n}(x)$ is normal to the tangent plane $\underline{n}(x) \cdot ds_x<u> = 0$ for any tangent vector u. Differentiating this equation gives

$$d\underline{n}_x< * > \cdot ds_x<u> + \underline{n}(x) \cdot d^2 s_x<u, *> = 0.$$

The $*$ in the above equation indicates that any tangent vector could be substituted here, the same tangent vector replacing both $*$s. This gives another form for the second fundamental form

$$II_x<u,v> = d^2 s_x<u,v> \cdot \underline{n}(x),$$

which is often more convenient to use as it does not involve differentiating the normal.

The equation

$$II_x<a,b> = \kappa \; I_x<a,b>$$

generically has two Eigenvectors p, q which are called **principal directions** and two corresponding Eigenvalues κ_p, κ_q the **principal curvatures**. Let $\underline{P} = ds_x<p>$, $\underline{Q} = ds_x<q>$ with p, q chosen so $\underline{P}$, $\underline{Q}$ are of unit length. These are the **principal directions on the surface** and they are independent of the parametrization and are orthogonal ([O'Neill] Theorem 5.2.5). In terms of the second fundamental form we have

$$\begin{aligned} II_x<p,p> &= -d\underline{n}_x<p> \cdot \underline{P} = \kappa_p, \\ II_x<q,q> &= -d\underline{n}_x<p> \cdot \underline{Q} = \kappa_q, \\ II_x<p,q> &= -d\underline{n}_x<p> \cdot \underline{Q} = -d\underline{n}_x<q> \cdot \underline{P} = 0. \end{aligned}$$

Two important measures of the shape of a surface at a point are the **Gaussian Curvature** $K = \kappa_p \kappa_q$ and the **Mean Curvature** $H = \frac{1}{2}(\kappa_p + \kappa_q)$. These values are independent of the parametrization ([O'Neill] lemma 3.2). The sign of the Gaussian curvature gives a basic classification of the shape of the surface at a point:

if $\kappa_p>0$, $\kappa_q>0$ or $\kappa_p<0$, $\kappa_q<0$ then $K>0$ and the surface at that point is said to be **elliptic**;

if $\kappa_p>0$, $\kappa_q<0$ or $\kappa_p<0$, $\kappa_q>0$ then $K<0$ and the surface at that point is said to be **hyperbolic**;

if one of κ_p or κ_q only is zero then the surface at the point is **parabolic**;

if both of κ_p and κ_q are zero then the surface at the point is **flat**.

On a generic surface you have regions of elliptical and hyperbolic points separated by lines of parabolic points (fig. 6.2).

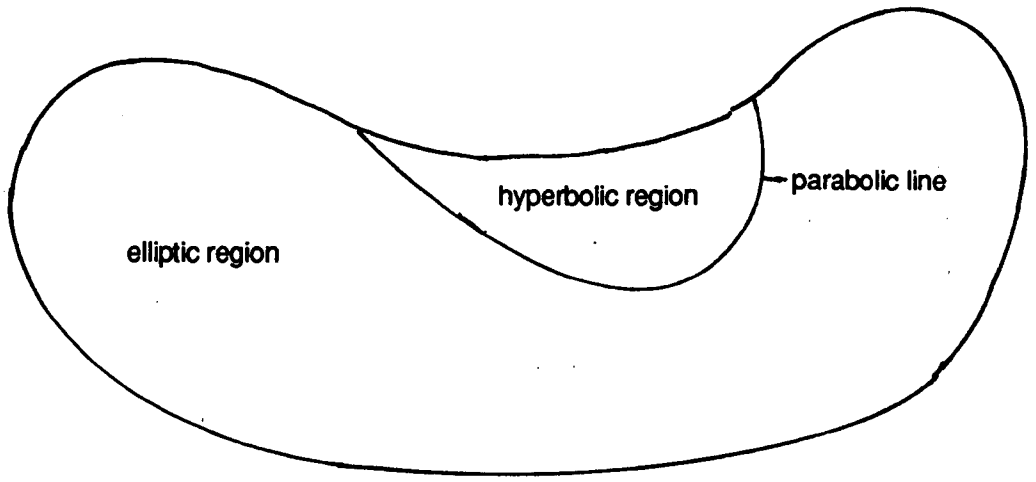


fig. 6.2) Elliptic, hyperbolic and parabolic regions of a typical surface

Rather than calculating the two principal curvatures we can calculate K and H directly from the first and second fundamental forms. First choose two linearly independent vectors u, v and let

$$E = I_x<u,v>, \qquad F = I_x<u,v>, \qquad G = I_x<v,v>,$$
$$l = II_x<u,u>, \qquad m = II_x<u,v>, \qquad n = II_x<v,v>.$$

A simple calculation shows that Gaussian and mean curvatures are

$$K = \frac{ln - m^2}{EG - F^2}, \qquad H = \frac{En + Gl - 2Fm}{EG - F^2}.$$

§6.3 Contact with spheres and the distance squared function

A standard way of classifying points on a surface is to examine the contact of the surface with a sphere. This is equivalent to studying the singularities of the **distance squared function**

$$V : \mathbf{R}^2 \times \mathbf{R}^3 \to \mathbf{R}$$
$$V(x, \underline{c}) = \tfrac{1}{2}(s(x) - \underline{c}) \cdot (s(x) - \underline{c}).$$

This gives half the square of the distance from a point on the surface $s(x)$, to a point $\underline{c}$ in space. See [Porteous–2].

The distance squared function is related to contact with spheres. The map $F_{\underline{c},\rho} : \mathbf{R}^3 \to \mathbf{R}$ defined by

$$F_{\underline{c},\rho}(\underline{S}) = (\underline{S} - \underline{c}) \cdot (\underline{S} - \underline{c}) - \rho^2,$$

defines a sphere $F^{-1}_{\underline{c},\rho}(0)$ centre $\underline{c}$ radius ρ. To find the contact of the surface with the sphere we examine the function $F_{\underline{c},\rho} \circ s : \mathbf{R}^2 \to \mathbf{R}$ which is our function $V|_{\underline{c}}$ after

adding and multiplying by constants. If we allow smooth changes of coordinates in the source we find that $V|_{\underline{c}}$ is right equivalent to one of the following forms:

$$\begin{aligned}
&A_0 : \ x_1 \pm x_2^2, \\
&A_1 : \pm x_1^2 \pm x_2^2, \\
&A_2 : \ x_1^3 \pm x_2^2, \\
&A_3 : \pm x_1^4 \pm x_2^2, \\
&A_4 : \ x_1^5 \pm x_2^2, \\
&D_4^+ : x_1^2 x_2 + x_2^3, \\
&D_4^- : x_1^2 x_2 - x_2^3.
\end{aligned}$$

This list is complete for a generic surface, a result of Looijenga – see [Wall]. In a generic one parameter family we can also have:

$$\begin{aligned}
&A_5 : \pm x_1^6 \pm x_2^2, \\
&D_5 : x_1^2 x_2 + x_2^4.
\end{aligned}$$

For example if $s(0)$ is a generic point on the surface and if $\underline{c}$ is a generic point a distance ρ along the normal line then a parametrization of the surface can be found so that

$$V|_{\underline{c}}(x) = \alpha(x_1^2 \pm x_2^2) \pm \beta,$$

for some constants α, β. We say that the sphere centre $\underline{c}$ radius ρ has A_1 **contact with the surface**, alternatively we label $\underline{c}$ as an A_1 point or say $V|_{\underline{c}}$ has an A_1 singularity at $x = 0$.

For a generic point on the surface all points not lying on the normal line are A_0 points, all point on the normal, bar two, are A_1 points and the points distance $1/\kappa_p$, $1/\kappa_q$ along the normal are both A_2 points (fig. 6.3). The higher singularities $A_3, A_4, D_4, \ldots$ only occur for special points on the surface. We shall look at these after the focal surface has been introduced.

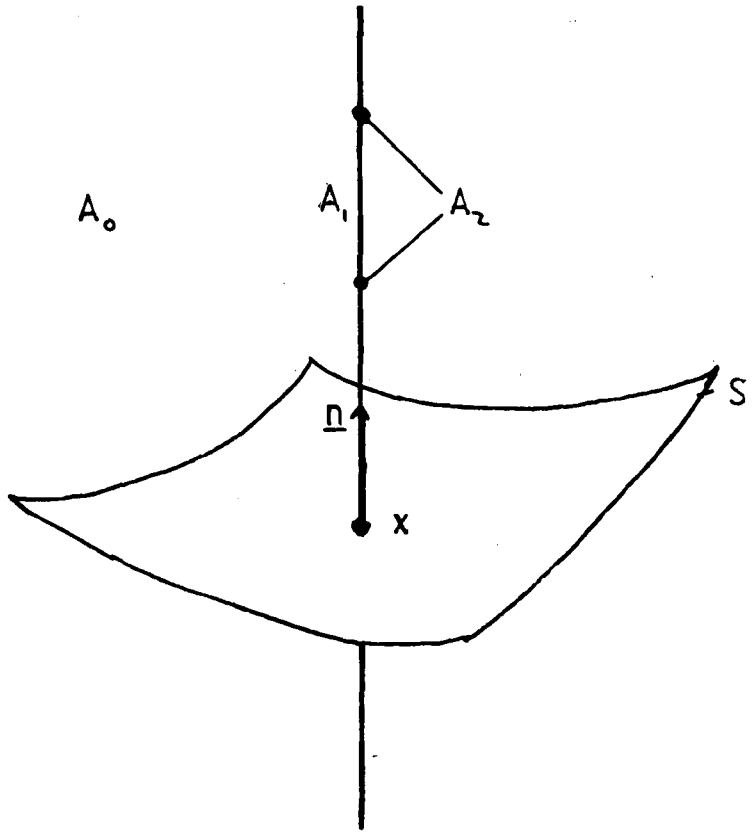


fig. 6.3) Labeling points in space according to singularities of distance squared map

§6.4 Focal surfaces and their singularities

The focal surface is the analogue of the evolute of a curve for surfaces in $\mathbf{R}^3$. Whereas the evolute is the locus of centers of circles with $A_{\geq 2}$ contact, the **focal surface** is the set of centers of spheres with A_2 or higher contact with the surface. We can also think of the focal surface as the envelope of the normal lines to the surface.

As there are two principal curvatures for most points on the surface there will be two A_2 points on the normal line and hence two sheets of the focal surface. These sheets will only come together at isolated points call umbilics.

Most points on the focal surface are A_2 points and at such point the focal surface is smooth. There can be lines of A_3 points, along such lines the focal surface has a cuspidal edge called a **rib**. On the surface S we have a corresponding **ridge line**. Along this line the principal curvature has a turning point $d\kappa_p<p> = 0$, so ridges are lines of "highest" or "lowest" principal curvature a feature which maybe visually evident. Isolated points on the ribs may be A_4 or **swallowtail** points. Here the

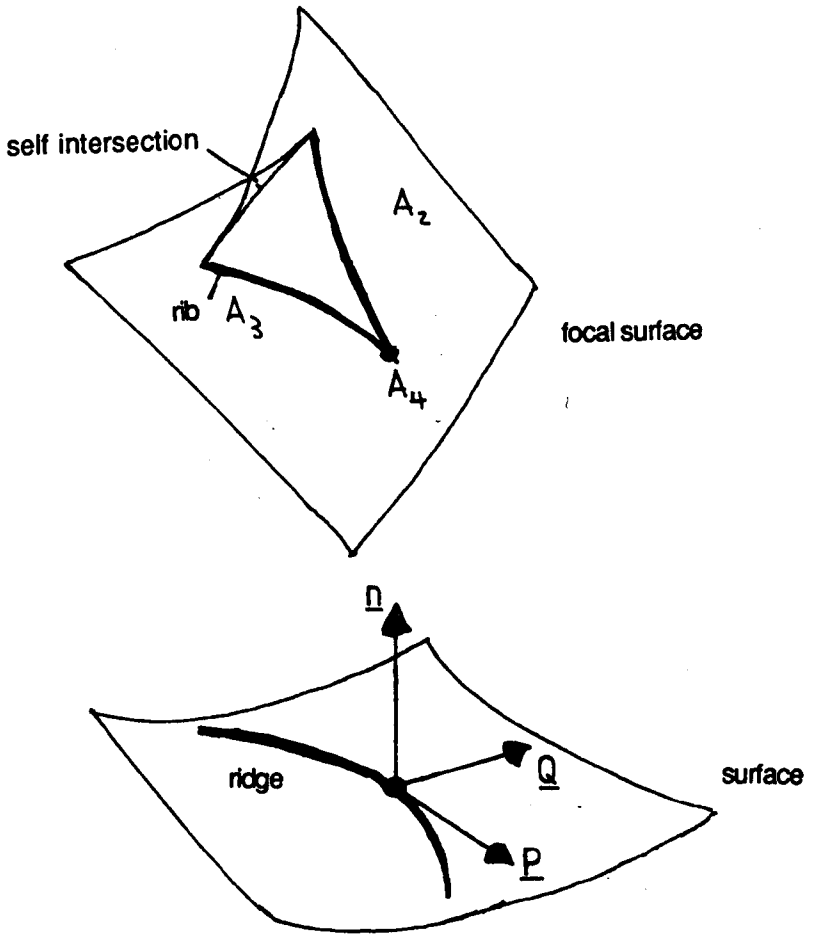


fig. 6.4) A swallowtail point on a focal surface

focal surface has a swallowtail singularity (fig 6.4). The line where the focal surface intersects itself is not particularly significant as the two sheets of the intersection come from different parts of the surface and do not even have the same principal curvatures. Hence there are two different spheres with A_2 contact.

When the two principal curvatures are equal, which for a generic surface only happens at isolated points, we have an **Umbilic point** lying on the surface and an **Umbilic centre** on the focal surface. Here the two sheets of the focal surface come together. These may be of two main types either **elliptic**, D_4^- (fig. 6.5a), or **hyperbolic**, D_4^+, (fig 6.5b). In the **parabolic-umbilic**, D_5 transition the type of the umbilic changes from being elliptic to hyperbolic. Through a hyperbolic umbilic centre there is a line of A_3 points which passes smoothly from one sheet to the other (prop. 8.12). There are three such lines through a elliptic umbilic centre. The pictures shown are the standard models of umbilics which are symmetrical. All the points on each surface have the same sign of Gaussian curvature and there are

no parabolic points. In chapter 8 we will look at generic examples of umbilics which are not symmetrical and for which have parabolic lines on them.

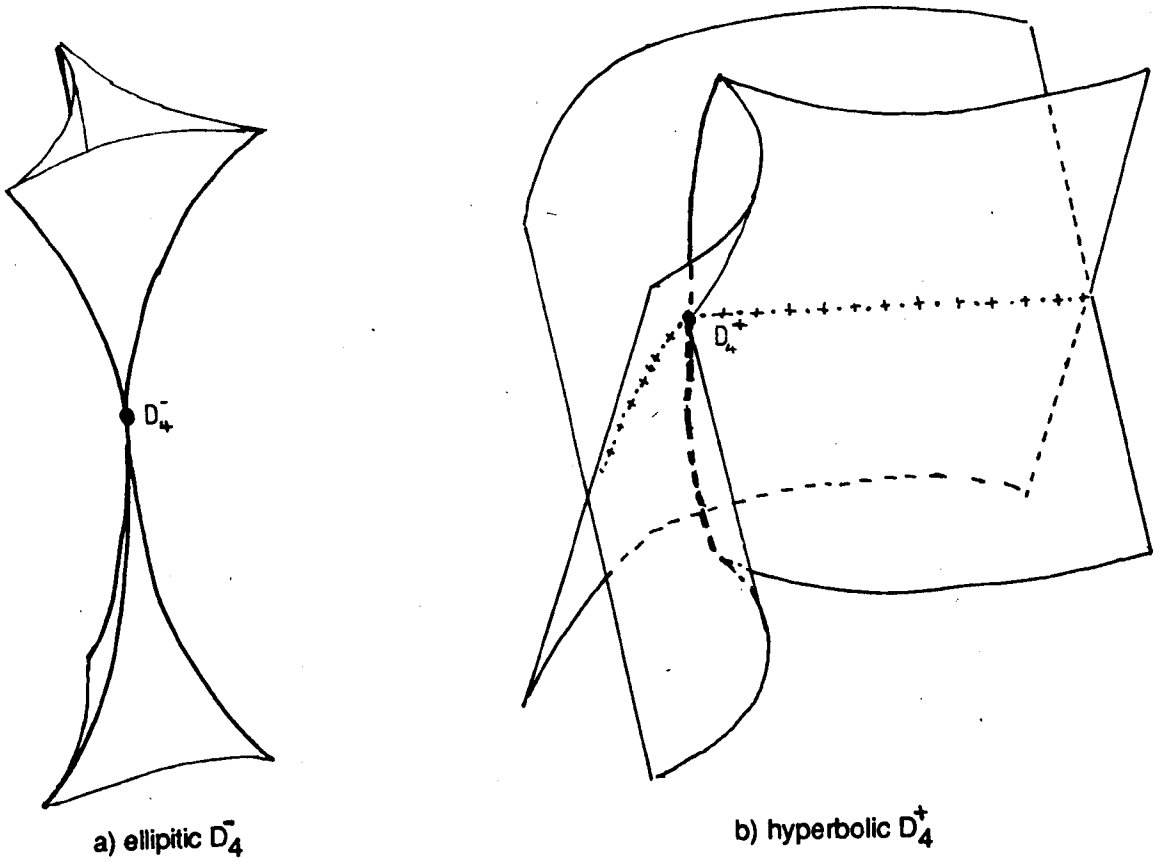


fig. 6.5) Umbilic focal surfaces (standard model)

Chapter 7: Calculations on focal surfaces

For many features of the surface (parabolic lines, ridges, parabolic lines on the focal surface) we can use the following technique for finding them: first find a function $f : \mathbf{R}^2 \to \mathbf{R}$, from the parameter space to the reals, such that $f^{-1}(0)$ (generically a set of lines) corresponds to the feature; once we know this function we can use a zero following algorithm to calculate the feature for an actual surface. For example to find the parabolic lines on a surface we use the Gaussian curvature as our function. Here we are interested in two different features the cuspidal edges (ribs) and the parabolic lines of the focal surface. The parabolic lines have only recently been considered. First theoretically as the singularities of folding maps [Bruce-Wilkinson] and now here where we present explicit expressions for calculating there positions. To study parabolic lines we have had to calculate the Gaussian curvature of the focal surface (§7.2, §7.3) and this has lead to an interesting discovery about cuspidal edges. Namely that the Gaussian curvature generically changes sign, through infinity, as we cross a cuspidal edge of a focal surface (Theorem 7.8).

Let $s : \mathbf{R}^2 \to \mathbf{R}^3$ be a parametrization of a surface and $f : \mathbf{R}^2 \to \mathbf{R}^3$ the corresponding parametrization of one sheet of the focal surface.

§7.1 The relationship between the derivative of the distance squared function and its singularities

Let $V : \mathbf{R}^2 \times \mathbf{R}^3 \to \mathbf{R}$ be the distance squared function introduced in chapter 6,

$$V(x, \underline{c}) = \tfrac{1}{2}(s(x) - \underline{c}) \cdot (s(x) - \underline{c}),$$

where $x = (x_1, x_2)$ is in the parameter space and $\underline{c}$ is a point in $\mathbf{R}^3$. Differentiating with respect to x we have

$$\begin{aligned}
dV(x,\underline{c})<u> &= ds(x)<u> \cdot (s(x) - \underline{c}),\\
d^2V<uv> &= d^2s<uv> \cdot (s - \underline{c}) + ds<u> \cdot ds<v>,\\
d^3V<uvw> &= d^3s<uvw> \cdot (s - \underline{c}) + d^2s<wu> \cdot ds<v>,\\
&\quad + d^2s<uv> \cdot ds<w> + d^2s<vw> \cdot ds<u>,\\
d^4V< > &= d^4s< > \cdot (s - \underline{c}) + 4d^3s< > \cdot ds< > + 3d^2s< > \cdot d^2s< >,\\
d^5V< > &= d^5s< > \cdot (s - \underline{c}) + 5d^4s< > \cdot ds< > + 10d^3s< > \cdot d^2s< >.
\end{aligned}$$

In general we assume that these maps are evaluated at a known point $(x, \underline{c})$ so we drop these parameters from the expressions. Also unless a specific tangent vector is

used we can either replace each occurrence by $*$ or by a space to indicate any tangent vector may be used here. For example the following are equivalent statements

$$dV(x,\underline{c})<u> = 0 \; \forall \; u \in \mathbf{R}^2,$$
$$dV<u> = 0 \; \forall \; u \in \mathbf{R}^2,$$
$$dV< > = 0,$$
$$dV<*> = 0.$$

In the fourth and fifth derivatives above we have abbreviated the dot product terms. When evaluating on a set of vectors the sum of various permutations should be used. The permutations are chosen so that each half of the dot product is evaluated on a unique sub-set of the vectors. For example when evaluating $4d^3s< > \cdot ds< >$ on the vectors u, v, w, z we calculate

$$d^3s<uvw> \cdot ds<z> \quad + d^3s<vwz> \cdot ds<u>$$
$$+ d^3s<wzu> \cdot ds<v> \quad + d^3s<zuv> \cdot ds<w>.$$

Note we have not included $d^3s<vuw> \cdot ds<z>$ as this is the same as $d^3s<uvw> \cdot ds<z>$

There is a link between these derivatives and the singularities of the distance squared function which was first explained in [Porteous–1] [Porteous–2]. We now summarize some of the results from [Porteous–2].

Lemma 7.1

First define the following conditions:

$dV(x,\underline{c})<*> = 0$ (7.1)

$d^2V(x,\underline{c})<p*> = 0$ for some non zero vector p (7.2)

$d^3V<p^3> = 0$ for p as in (7.2) (7.3)

The equations (7.4)

$$d^3V<p^2*> + d^2V<r*> = 0$$
$$d^4V<p^4> + 6d^3V<p^2r> + 3d^2V<r^2> = 0$$

have a simultaneous solution for some non zero vector r and p as in (7.2)

$d^5V<p^5> + 10d^4V<p^3r> + 15d^3V<pr^2> = 0$ p as in (7.2), r as in (7.4) (7.5)

$d^2V<**> = 0$ (7.6)

$d^3V<p^2*> = 0$ for some non zero vector p (7.7)

$d^3V<p**> \neq 0$ for p as in (7.7) (7.8)

$d^4V<p^4> = 0$ for p as in (7.7) (7.9)

Note if (7.1) holds then c must lie on the normal line, furthermore if (7.2) also holds then c is a focal point and p is the corresponding principal direction. The other conditions are all evaluated at the point $(x, \underline{c})$. The conditions for the various singularities of the distance squared map are:

for A_1 points (7.1) but not (7.2) holds;
for A_2 points (7.1), (7.2) but not (7.3) or (7.6) hold;
for A_3 points (7.1), (7.2), (7.3) but not (7.4) or (7.6) hold;
for A_4 points (7.1), (7.2), (7.3), (7.4) but not (7.5) or (7.6) hold.;
for D_4 points (7.1), (7.6) but not (7.7) hold.
for D_5 points (7.1), (7.6), (7.7) but not (7.8) or (7.9) hold;

Corollary 7.2

The ribs and ridges (A_3 or higher points) are given by the set of points in the parameter space

$$R = \left\{ \begin{array}{ll} x \in \mathbf{R}^2 & \text{such that } d^3V(x,\underline{c})<p^3> = 0, \text{ where } p \text{ is a principal direction} \\ & \text{and } \underline{c} \text{ is the corresponding focal point.} \end{array} \right.$$

The ridges (on the surface) are given by $s(R)$ and the ribs (on the focal surface) are $f(R)$. To find A_4 or higher points we look for the intersection of R with

$$T = \left\{ \begin{array}{ll} x \in \mathbf{R}^2 & \text{such that } 3d^3V<p^2q>^2 - d^4V<p^4>d^2V<q^2> = 0, \\ & \text{where } p \text{ is a principal direction,} \end{array} \right.$$

$f(T \cap R)$ gives the swallowtail points on the focal surface and $s(T \cap R)$ gives the corresponding points on the surface.

Proof

An A_4 points must lie on a rib and condition (7.4) must also hold. Let $r = \alpha p + \beta q$. Condition (7.4) implies

$$d^3V<p^2*> + \beta d^2V<q*> = 0 \qquad (7.10)$$

$$d^4V<p^4> + 6\beta d^3V<p^2q> + 3\beta^2 d^2V<q^2> = 0 \qquad (7.11)$$

The first of these equations implies that the derivative of the second with respect to β is zero. Hence the quadratic in β (7.11) has repeated roots and a zero discriminant. Taking the discriminant we have

$$(3d^3V<p^2q>)^2 - 3d^2V<q^2>\ d^4V<p^4> = 0.$$

§7.2 Calculating the Gaussian curvature of the focal surface

Now we begin the new work by looking at the derivatives of one sheet of the focal surface f and of the principal direction $\underline{P}$ associated with this sheet. This will lead to an expression for the Gaussian curvature. Previously when differentiating we have kept $\underline{c}$, the point on the focal surface, fixed. If we now allow $\underline{c}$ to vary but restrict its movement to the focal surface we can obtain expressions for $d\underline{c}< > = df< >$ and $d^2\underline{c}< > = d^2f< >$ and hence information about the focal surface.

Let $C : \mathbf{R} \to \mathbf{R}^2$ be a curve in the parameter space, $s \circ C : \mathbf{R} \to \mathbf{R}^3$ is a curve on the surface and $f \circ C : \mathbf{R} \to \mathbf{R}^3$ is the corresponding curve on the focal surface. We have two vector fields $p, q : \mathbf{R}^2 \to \mathbf{R}^2$ for the principal directions in the parameter space where p corresponds to the sheet of the focal surface given by f. We also have two vector fields $\underline{P}, \underline{Q} : \mathbf{R}^2 \to \mathbf{R}^3$ which are the actual principal

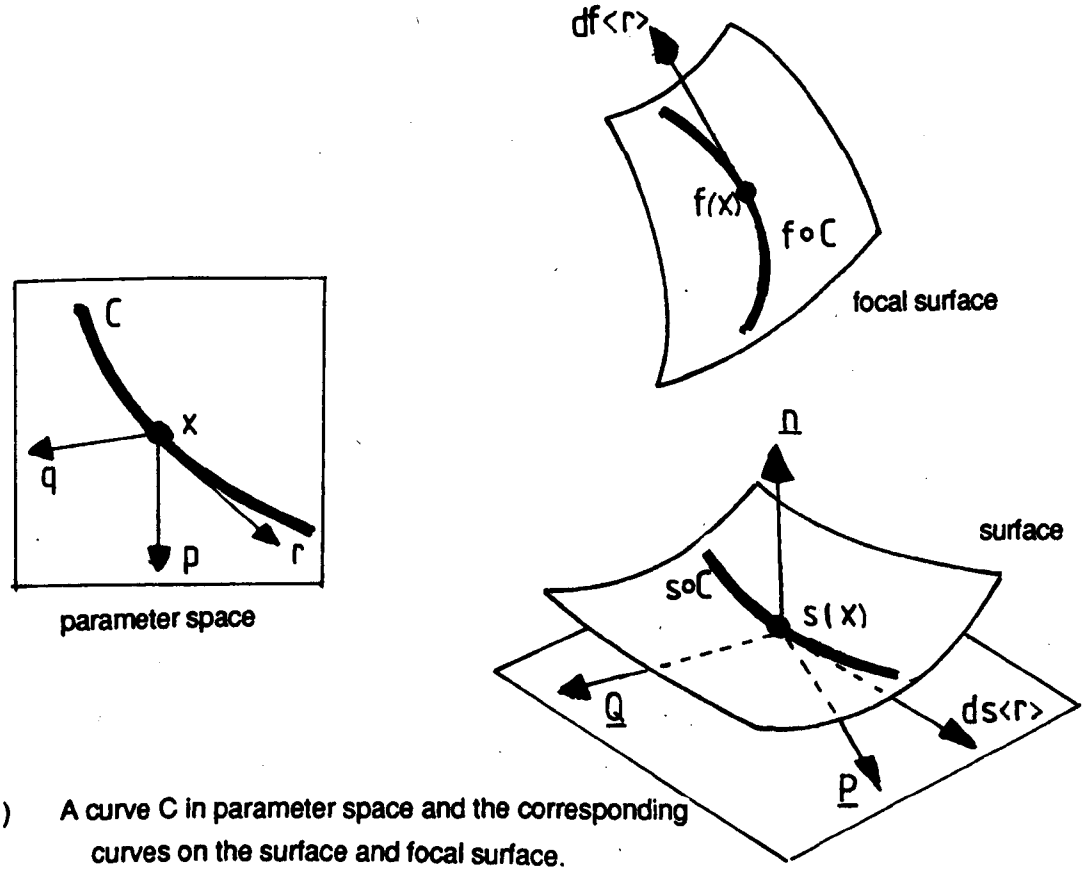


fig. 7.1) A curve C in parameter space and the corresponding curves on the surface and focal surface.

directions on the surface (fig. 7.1). They are defined by $\underline{P}(x) = ds(x)<p(x)>$ and $\underline{Q}(x) = ds(x)<q(x)>$. We choose p, q such that $\underline{P}$ and $\underline{Q}$ are unit length. Provided we keep away from umbilic points these vector fields are continuous. If $x = C(0)$

and $r = dC(0) \in \mathbf{R}^2$ then

$$\begin{aligned} ds_x<r> &= d(s \circ C)(0) \in \mathbf{R}^3 \\ df_x<r> &= d(f \circ C)(0) \in \mathbf{R}^3 \\ dp_x<r> &= d(p \circ C)(0) \in \mathbf{R}^2 \\ d\underline{P}_x<r> &= d(\underline{P} \circ C)(0) \in \mathbf{R}^3. \end{aligned}$$

We can now differentiate the equation

$$dV(x, f(x))<*> = ds(x)<*> \cdot (s(x) - f(x)) = 0$$

along the curve C to get

$$\begin{aligned} d^2s_x<r*> \cdot (s(x) - f(x)) + ds_x<*> \cdot ds_x<r> - ds_x<*> \cdot df_x<r> \\ = d^2V(x, f(x))<r*> - ds_x<*> \cdot df_x<r> \\ = 0. \end{aligned} \tag{7.12}$$

Replacing $*$ by p gives

$$ds_x<p> \cdot df_x<r> = \underline{P}(x) \cdot df_x<r> = 0.$$

From this equation we obtain our first piece of information about the focal surface.

Lemma 7.3

The principal direction $\underline{P}(x)$ is normal to the focal surface.

Proof

From [Porteous-3] p85. As $\underline{P}(x) \cdot df_x<r> = 0$ for all vectors $r \in \mathbf{R}^2$ the vector $\underline{P}(x)$ is normal to the tangent plane to the focal surface and hence to the focal surface itself.

Henceforth we assume that derivatives are evaluated at x, {$(x, f(x))$ in the case of the distance squared function}, and we drop subscripts. To find out what the tangent vector $df<r>$ is we can use equation (7.12) again, this time replacing $*$ by q to give

$$df<r> \cdot \underline{Q} = d^2V<rq>. \tag{7.13}$$

Differentiating $d^2V<p*> = 0$ along C gives

$$d^3V<pr*> + d^2V<dp<r>*> - d^2s<p*> \cdot df<r> = 0. \tag{7.14}$$

By replacing $*$ by p we have

$$d^3V<p^2r> = d^2s<p^2> \cdot df<r>. \tag{7.15}$$

Now let $df<r> = a\ \underline{Q} + b\ \underline{n}$; we find from (7.15) that $a = d^2V<qr>$. Expanding (7.15) gives

$$d^3V<p^2r> = a\ d^2s<p^2> \cdot \underline{Q} + b\ d^2s<p^2> \cdot \underline{n}$$

so

$$b = \frac{d^3V<p^2r> - d^2V<qr>\ d^2S<p^2> \cdot \underline{Q}}{d^2s<p^2> \cdot \underline{n}}.$$

Lemma 7.4

The tangent vector $df<r>$ to the curve $f \circ C$ where $r = \frac{dC}{dt}$ is

$$d^2V<qr>\underline{Q} + \left(\frac{d^3V<p^2r> - d^2V<qr>\ d^2s<p^2> \cdot \underline{Q}}{\kappa_p}\right)\underline{n}.$$

Proof

From the above noting that the principal curvature $\kappa_p = d^2s<p^2> \cdot \underline{n}$.

We can now find the **first fundamental form of the focal surface,**

$$I_f<u,v> = df<u> \cdot df<v>.$$

To find the second fundamental form (and hence the Gaussian curvature) we can either find the second derivative of the focal surface $d^2f< >$ or the first derivative of the normal to the focal surface. As the normal is $\underline{P}$ it will be easier to do the latter.

Now $\underline{P} = ds<p>$ so

$$d\underline{P}<r> = d^2s<pr> + ds<dp<r>>$$

which means we have to find $dp<r>$. We have chosen p such that $\underline{P}$ is of unit length i.e. $\underline{P} \cdot \underline{P} = 1$. Differentiating this gives

$$d\underline{P}<r> \cdot \underline{P} = d^2s<pr> \cdot \underline{P} + d^2s<dp<r>> \cdot \underline{P} = 0 \tag{7.16}$$

Another equation for $dp<r>$ is obtained by replacing $*$ by q in (7.14),

$$d^3V<pqr> + d^2V<q\ dp<r>> - d^2s<pq> \cdot df<r> = 0. \tag{7.17}$$

Letting $dp<r> = a\ p + b\ q$ equation (7.16) becomes

$$\begin{aligned} 0 &= d^2s<pr> \cdot \underline{P} + a\ ds<p> \cdot \underline{P} + b\ ds<q> \cdot \underline{P} \\ &= d^2s<pr> \cdot \underline{P} + a, \end{aligned}$$

as $ds<p> = \underline{P}$ and $ds<q> = \underline{Q}$ are orthogonal unit vectors. Also equation (7.17) reduces to

$$d^3V<pqr> + b\ d^2V<q^2> - d^2s<pq> \cdot df<r> = 0.$$

We can now find a, b and hence $dp<r>$,

$$dp<r> = (-d^2s<pr> \cdot \underline{P})p + \left(\frac{d^2s<pq> \cdot df<r> - d^3V<pqr>}{d^2V<q^2>} \right) q. \qquad (7.18)$$

Lemma 7.5

The derivative of the normal to the focal surface is

$$d\underline{P}<r> = d^2s<pr> - (d^2s<pr> \cdot \underline{P})\underline{P} + \left(\frac{d^2s<pq> \cdot df<r> - d^3V<pqr>}{d^2V<q^2>} \right) \underline{Q}.$$

Definition 7.6

The **second fundamental form of the focal surface** is

$$II_f<u, v> = -df<u> \cdot d\underline{P}<v>.$$

To find the Gaussian curvature first choose two linearly independent vectors u, v and let

$$\begin{aligned} E &= I_f<u, v>, \qquad F = I_f<u, v>, \qquad G = I_f<v, v>, \\ l &= II_f<u, u>, \qquad m = II_f<u, v>, \qquad n = II_f<v, v> \end{aligned}$$

and the **Gaussian curvature of the focal surface** is

$$K_f = \frac{ln - m^2}{EG - F^2}.$$

§7.3 A simplified form for the Gaussian curvature of the focal surface

The previous expression can be greatly simplified by choosing $u = p$ and $v = q$. For the tangents we have

$$\begin{aligned} df<p> &= \frac{1}{\kappa_p} d^3V<p^3> \underline{n}, \\ df<q> &= d^2V<q^2> \underline{Q} + \left(\frac{d^3V<p^2q> - d^2V<q^2>\ d^2s<p^2> \cdot \underline{Q}}{\kappa_p} \right) \underline{n}. \end{aligned}$$

We observe that when $d^3V<p^3> = 0$ the map $df : \mathbf{R}^2 \to \mathbf{R}^3$ is a fold map. This ties in with this also being the condition for a cuspidal edge (A_3 point) on the focal surface.

Before moving on to find the curvature of the focal surface we first have a lemma about A_4 points.

Lemma 7.7

If the tangent to the ridge is in the principal direction then we have an A_4 singularity.

Proof

This is also stated without proof in [Porteous–2] p382. If r is the tangent to the ridge then differentiating $d^3V<p^3> = 0$ in the r direction gives

$$d^4V<p^3r> + 3d^3V<p^2\ dp<r>> - d^3s<p^3> \cdot df<r> = 0.$$

If $r = p$, the principal direction, this reduces to

$$d^4V<p^4> + 3d^3V<p^2 dp<p>> = 0.$$

From equation (7.17) we also have

$$d^3V<p^2*> + d^2V<dp<p>*> = 0.$$

These two equations imply condition (7.4) of lemma 7.1 with $r = dp<p>$ so we have an A_4 point.

We now return to the task of calculating the Gaussian curvature of the focal surface and find the derivatives of $\underline{P}$. Recalling that $d^2s<pq> \cdot \underline{n} = 0$, $d^2s<p^2> \cdot \underline{n} = \kappa_p$ the expressions for $d\underline{P}<p>$, $d\underline{P}<q>$ simplify to

$$d\underline{P}<p> = \left(d^2s<p^2> \cdot \underline{Q} - \frac{d^3V<p^2q>}{d^2V<q^2>} \right) \underline{Q} + \kappa_p\ \underline{n},$$

$$d\underline{P}<q> = \left(2d^2s<pq> \cdot \underline{Q} - \frac{d^3V<pq^2>}{d^2V<q^2>} \right) \underline{Q}.$$

Calculating the number E, F, G, l, m, n we have

$$E = (\tfrac{1}{\kappa_p} d^3V<p^3>)^2$$

$$F = \tfrac{1}{\kappa_p^2} d^3V<p^3> (d^3V<p^2q> - d^2V<q^2>\ d^2s<p^2> \cdot \underline{Q})$$

$$G = (d^2V<q^2>)^2 + \left(\frac{d^3V<p^2q> - d^2V<q^2>\ d^2s<p^2> \cdot \underline{Q}}{\kappa_p} \right)^2,$$

$$l = -d^3V<p^3>,$$

$$m = 0,$$

$$n = -2d^2V<q^2>\ d^2s<pq> \cdot \underline{Q} - d^3V<pq^2>.$$

We can now calculate the Gaussian curvature of the focal surface.

Theorem 7.8

The Gaussian curvature of the focal surface is given by

$$K_f = \frac{\overline{n}}{l}\frac{\kappa_p^2}{d^2V<q^2>}$$

where κ_p is the principal curvature of the surface S in the $\underline{P}$ direction and

$$l = -d^3V<p^3>, \qquad \overline{n} = -d\underline{P}<q> \cdot \underline{Q}.$$

(If $\kappa_p = 0$ the focal surface lies at infinity and if $d^2V<q^2> = 0$ we have an umbilic point.)

At a generic A_3 point $l = 0$ and $\overline{n}$, κ_p, $d^2V<q^2>$ are all non zero. The focal surface has a cuspidal edge with infinite Gaussian curvature, on either side of the edge the Gaussian curvature will be of opposite signs.

The focal surface has parabolic line, $K_f = 0$, when $\overline{n} = 0$ and l, κ_p, $d^2V<q^2>$ are all non zero. The curvatures on either side of the parabolic line will be of opposite signs.

Proof

Calculating the denominator of the Gaussian curvature, $K_f = (ln - m^2)/(EG - F^2)$, we find

$$\begin{aligned} EG - F^2 &= \left(\frac{d^3V<p^3>\, d^2V<q^2>}{\kappa_p}\right)^2 \\ &= l^2\left(\frac{d^2V<q^2>}{\kappa_p}\right)^2. \end{aligned}$$

Also by expressing $df<q>$ in terms of $\underline{P}$, $\underline{Q}$, $\underline{n}$ we have

$$\begin{aligned} n &= -df<q> \cdot d\underline{P}<q> \\ &= -(d^2V<q^2>\underline{Q} + \alpha\underline{n}) \cdot d\underline{P}<q> \\ &= -d^2V<q^2>\, d\underline{P}<q> \cdot \underline{Q} \\ &= d^2V<q^2>\overline{n}, \end{aligned}$$

where α is some real number. Substituting in the expression for the curvature gives the required result.

The map $l : \mathbf{R}^2 \to \mathbf{R}$ will have rank 0 if and only if both $dl<p>$ and $dl<q>$ are zero. From the proof of lemma 7.7 we see that $l = 0$, $dl<p> = 0$ implies that we have

an A_4 or higher singularity, which generically occurs at isolated points. Furthermore we also require $dl<q> = 0$ and this gives the non-generic A_4 singularity [Porteous–2] p382. It follows that at all points of a rib l will have rank 1, so its sign and the sign of K_f will change as we cross the rib. Likewise $\overline{n} : \mathbf{R}^2 \to \mathbf{R}$ generically has rank 0 only at isolated points and generically the curvature of the focal surface will change sign when we cross a parabolic line.

We note that in the standard model of a cusp edge, $f(x, y) = (x, y^2, y^3)$, the Gaussian curvatures are the same on either side of the edge (see example 1 below) but there is generically a sign change for cusp edges of focal surfaces. Now we look at the two conditions ($l = 0, \overline{n} = 0$) to see what they imply geometrically.

§7.4 Parabolic lines and cuspidal edges on focal surfaces

First consider a point $s(x)$ on a surface S for which $d\underline{n}<r> = 0$ for some vector r. If we let $r = a\ p + b\ q \neq 0$ we have

$$d\underline{n}<a\ p + b\ q> \cdot \underline{P} = -a\kappa_p = 0,$$
$$d\underline{n}<a\ p + b\ q> \cdot \underline{Q} = -b\kappa_q = 0$$

This implies that either $b = \kappa_p = 0$ and $r = a\ p$ or $a = \kappa_q = 0$ and $r = b\ q$ or that $\kappa_p = \kappa_q = 0$ which is a **flat point** (an umbilic with zero Gaussian Curvature). In all three cases we are at a parabolic point and r is a principal direction. Conversely if we are at a parabolic point, $\kappa_p = 0$, and p is a principal direction then $d\underline{n}<p> = 0$.

We now apply this to a point on a focal surface. Here $\underline{P}$ is a unit normal, so for a parabolic point we require that $d\underline{P}<r> = 0$ for some r. Therefore $d\underline{P}<r> \cdot \underline{n} = -\underline{P} \cdot d\underline{n}<r> = II_x<pr> = 0$ which implies $r = q$. Note II_x is the second fundamental form for the surface S at the point x. To be parabolic we also require that $d\underline{P}<r> \cdot \underline{Q} = \overline{n} = 0$, the same result as in theorem 7.8. Alternatively $\overline{n} = 0$ implies $d\underline{P}<q> = 0$. We note that the point is not flat as $d\underline{P}<p> \cdot \underline{Q}$ is generically non zero and at parabolic points $df<q>$ is a principal direction of the focal surface.

Intuitively we observe that when moving along a curve $s \circ C$, on the original surface S, in the $\underline{Q}$ direction the vector $\underline{P}$ remains stationary. Hence the normal to the focal surface does not change when moving along the curve $f \circ C$ in the $df<q>$

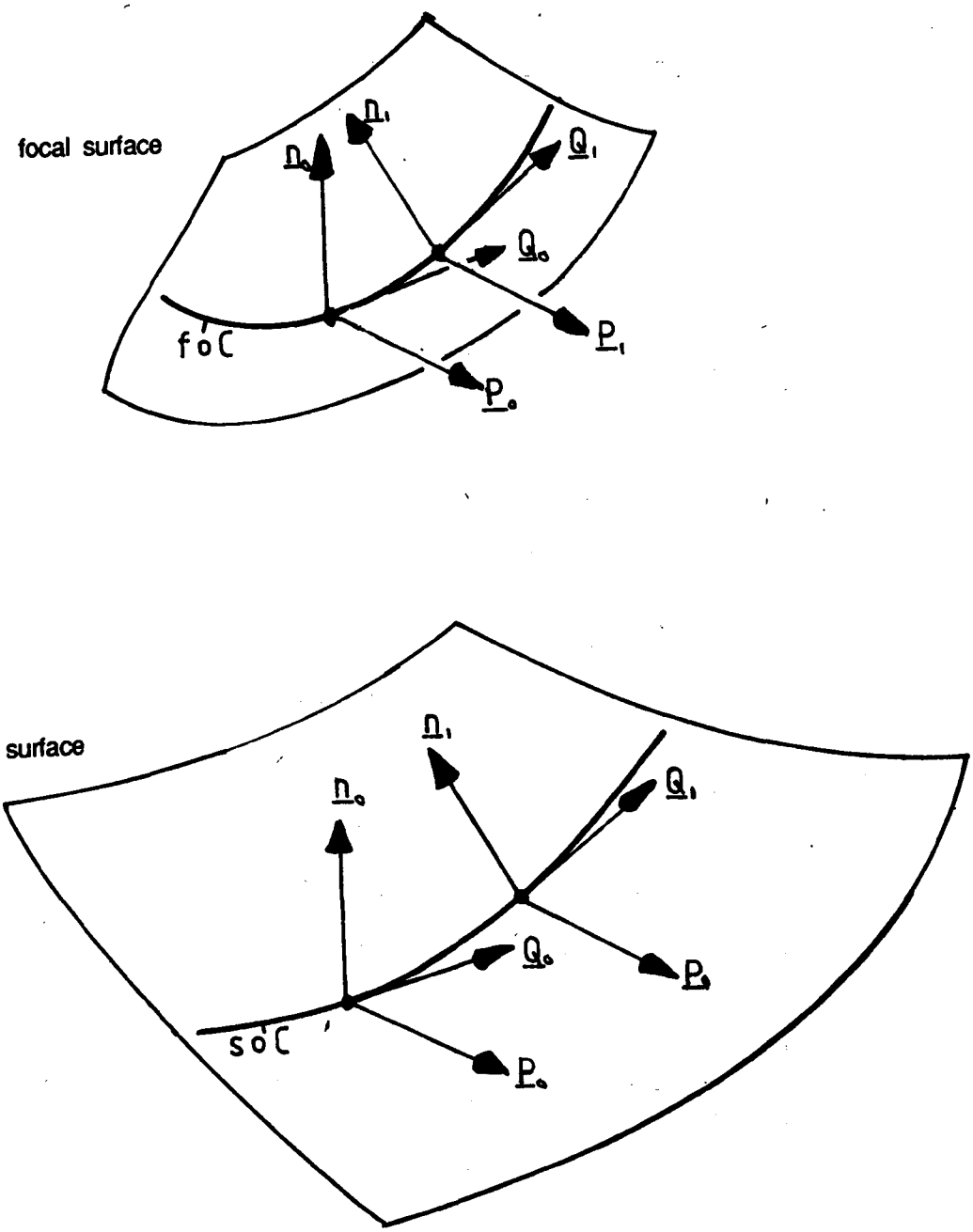


fig. 7.2) Change in $\underline{P}$, $\underline{Q}$ $\underline{n}$ frame at a parabolic point of the focal surface

direction (fig. 7.2). The significance of this on the surface is discussed in lemma 8.11.

Definition 7.9

For a parabolic line on the focal surface we can think about the corresponding curve on the original surface S. We shall call this curve a **sub-parabolic line**. Both of these curves can be defined in parameter space. If

$$T = \begin{cases} x \in \mathbf{R}^2 & \text{such that } d\underline{P}<q> \cdot \underline{Q} = 0, \\ & \text{where } p,\ q \text{ are principal directions,} \end{cases}$$

then the parabolic lines on the focal surface are given by $f(T)$ and the sub-parabolic lines are given by $s(T)$.

To analyse what happens near a cuspidal edge we define a basis $\underline{A}$, $\underline{B}$, $\underline{C}$ at every point of the edge with $\underline{A}$ along the edge, $\underline{B} = \underline{P}$ the limit of the normal to the

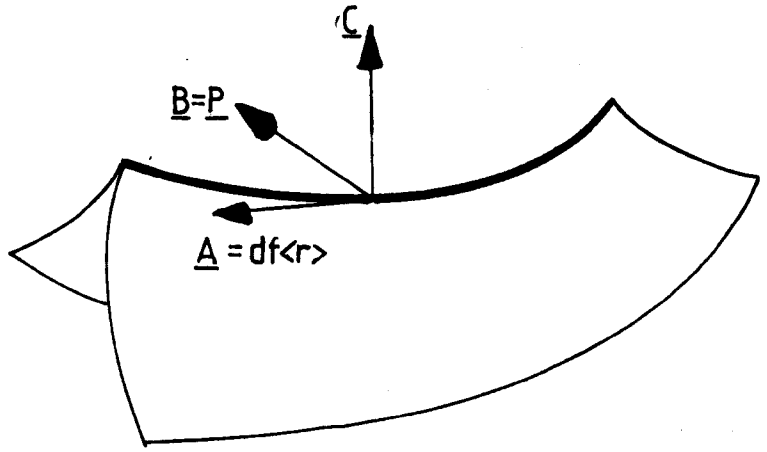


fig. 7.3) The frame A, B, C on a cuspidal edge

focal surface and $\underline{C} = \underline{B} \wedge \underline{A}$ (fig. 7.3). If $\gamma : \mathbf{R} \to \mathbf{R}^3$ is a unit speed parametrization of the cuspidal edge then $\underline{A} = \gamma'$ where primes denote differentiation along the edge. We can describe the shape of the cuspidal edge by three numbers:

$$\kappa = \underline{A}' \cdot \underline{B} = -\underline{B}' \cdot \underline{A}$$
$$g = \underline{A}' \cdot \underline{C} = -\underline{C}' \cdot \underline{A}$$
$$\tau = \underline{C}' \cdot \underline{B} = -\underline{B}' \cdot \underline{C}$$

These number are similar to those used to describe a curve on a smooth surface, see for example [Konederink] §6.1. If we think of A as the tangent to the curve and B as normal to the (cusp edge) surface then κ, g, τ are equivalent to the three quantities "normal curvature", "geodesic curvature" and "geodesic torsion" for a curve on a smooth surface. At a generic point on the cuspidal edge we would expect these three numbers to be non zero. The geometric significance of these numbers is illustrated below by a few special cases.

The colour pictures (7.5a)–c)*) are computer generated examples of some special cases and they have been calculated as the focal surfaces of some specially chosen surfaces. The colours are just used to give an indication of the curvature, rather than the specific values, and are coded as follows:

Orange - large positive Gaussian curvature,
Green - positive Gaussian curvature,

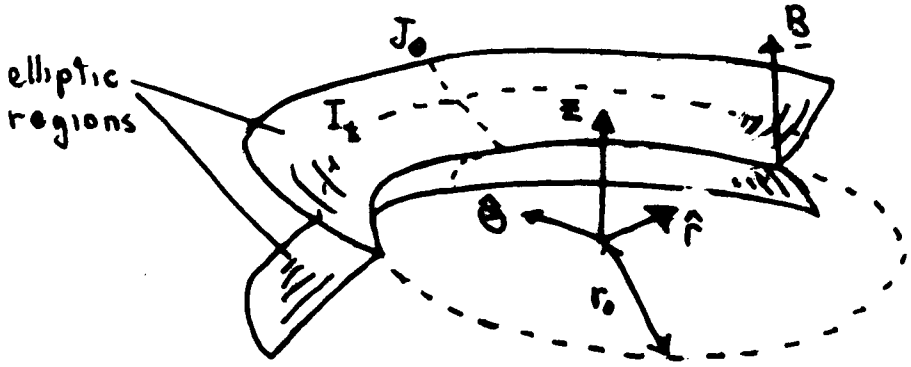


fig. 7.6) Overlay for fig 7.5a) showing elliptic region

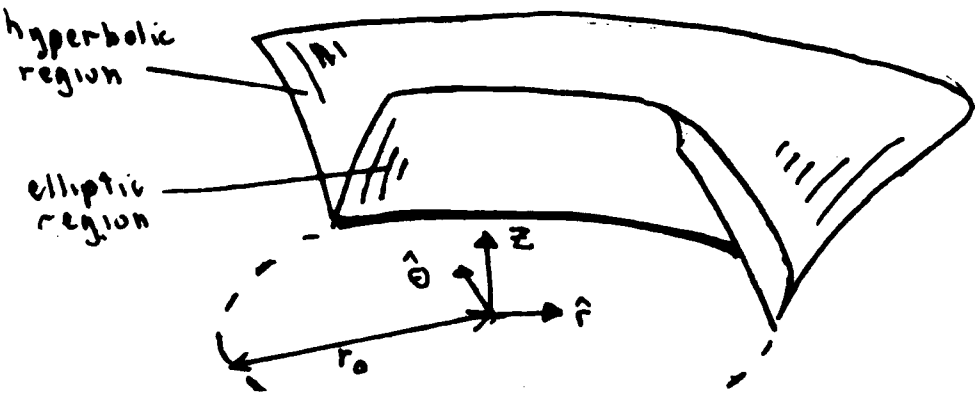


fig. 7.7) Overlay for fig 7.5b) showing elliptic and hyperbolic regions

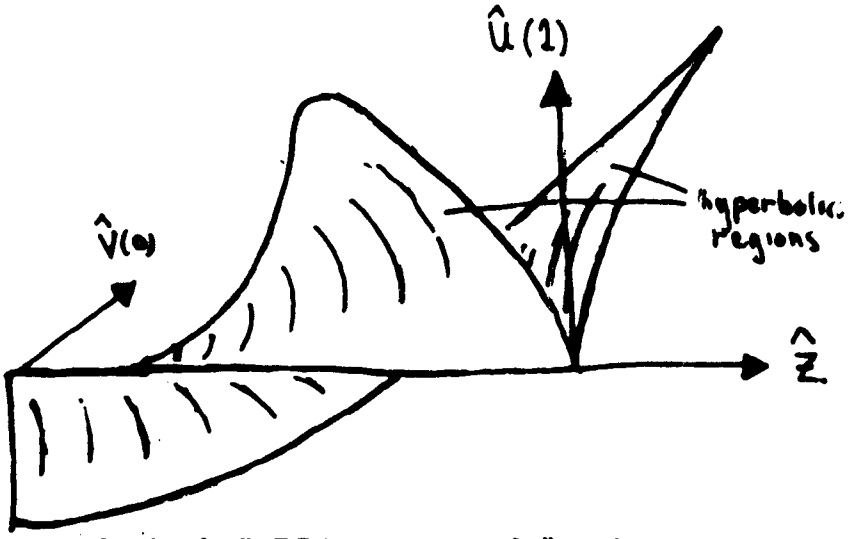


fig. 7.8) Overlay for fig 7.5c) showing hyperbolic region

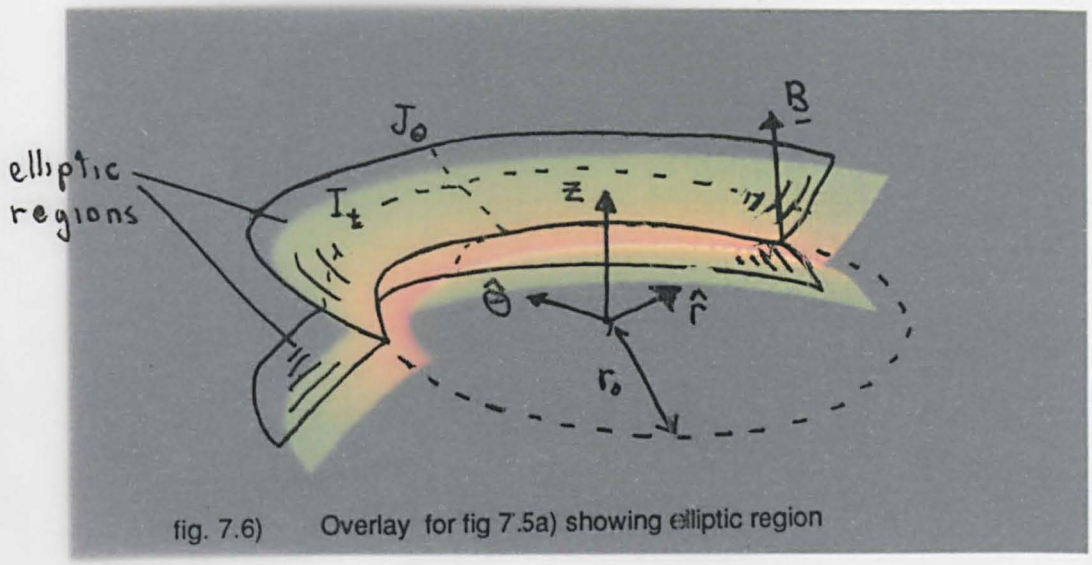


fig. 7.6) Overlay for fig 7.5a) showing elliptic region

a) $\kappa = \tau = 0$, g = constant.

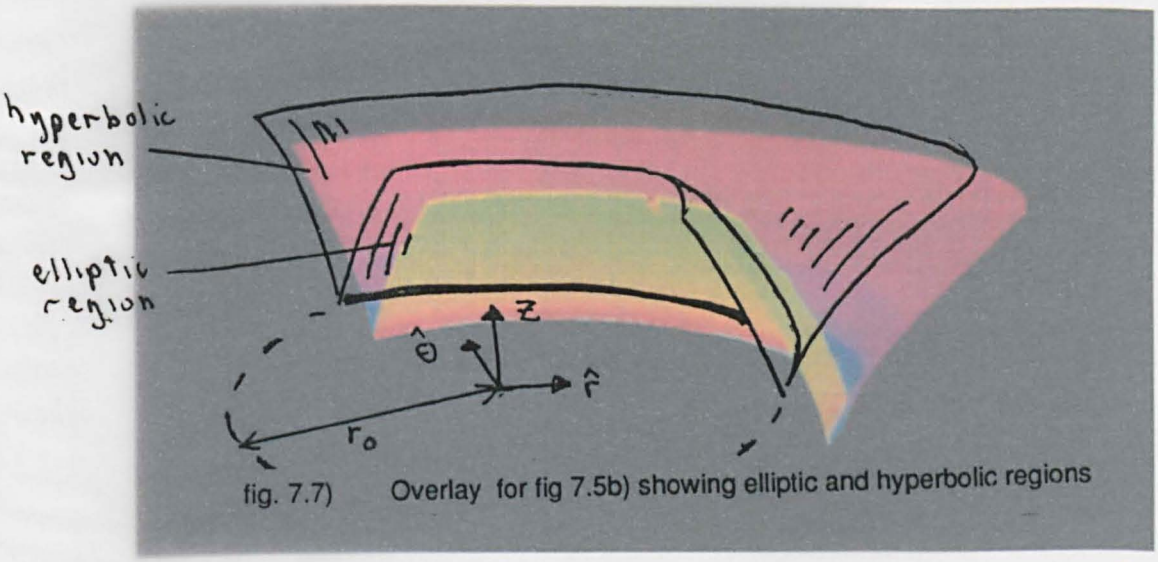


fig. 7.7) Overlay for fig 7.5b) showing elliptic and hyperbolic regions

b) $g = \tau = 0$, κ = constant.

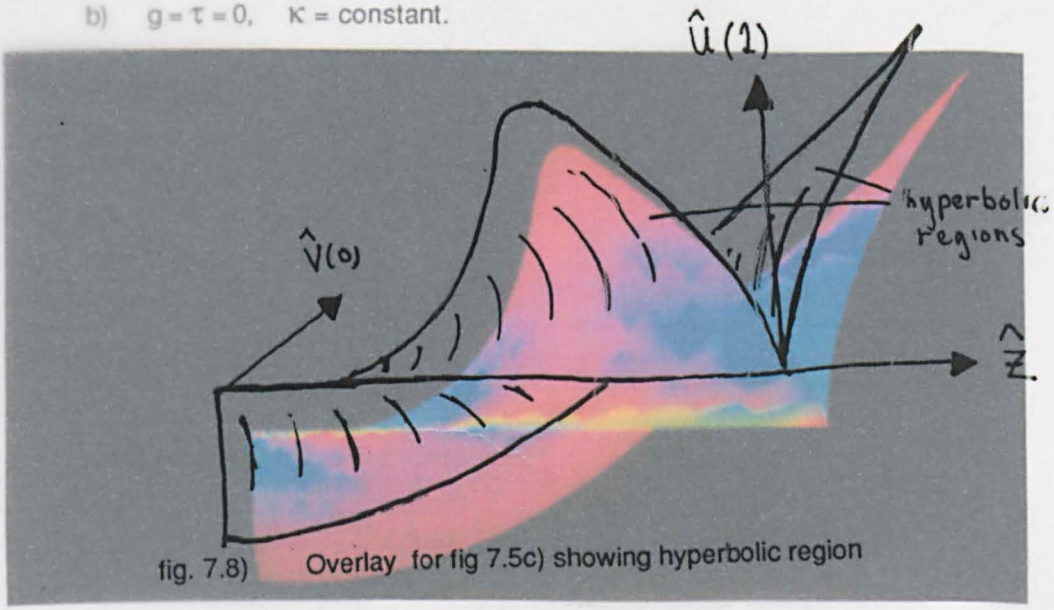
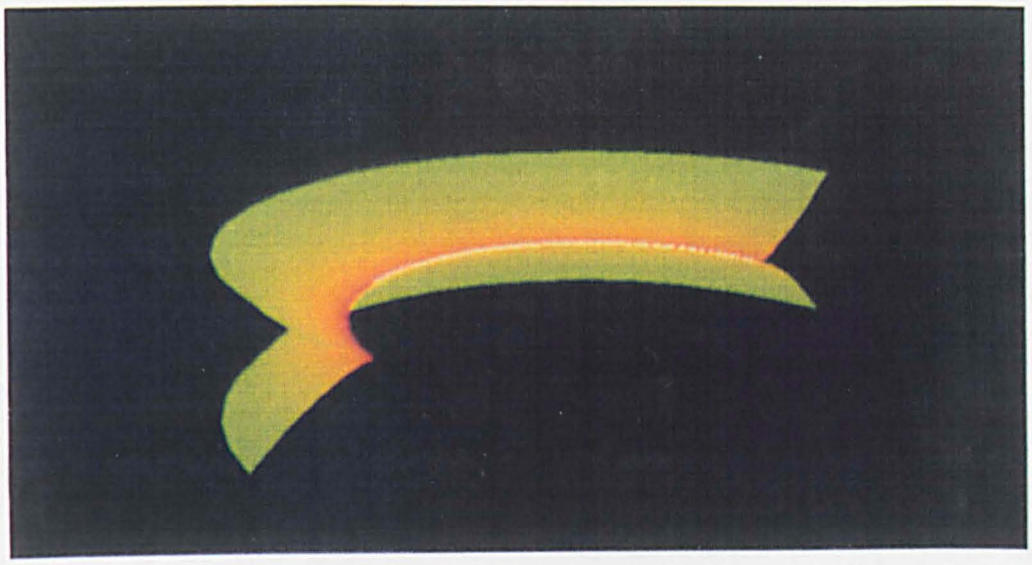


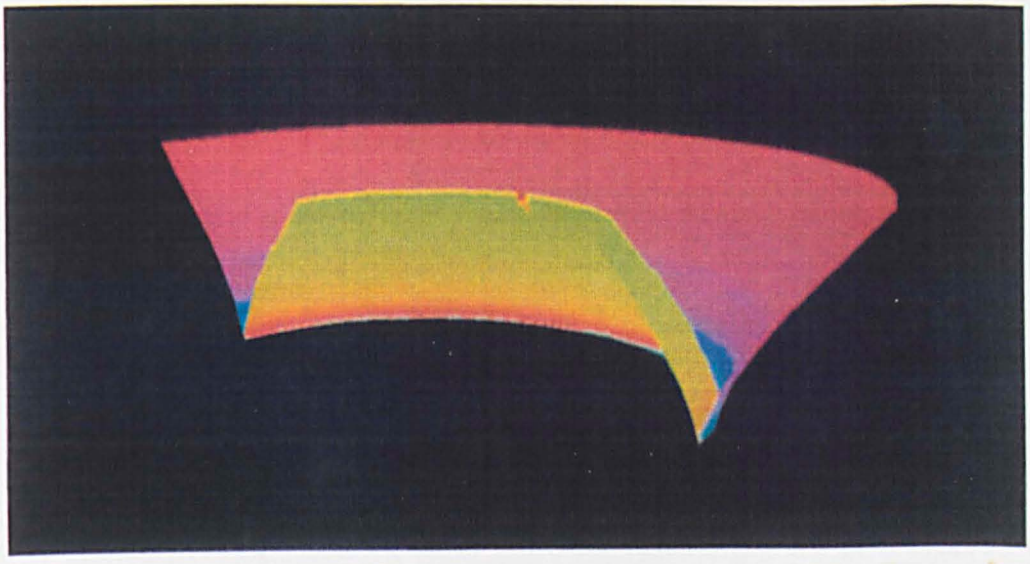
fig. 7.8) Overlay for fig 7.5c) showing hyperbolic region

c) $\kappa = g = 0$, τ = constant.

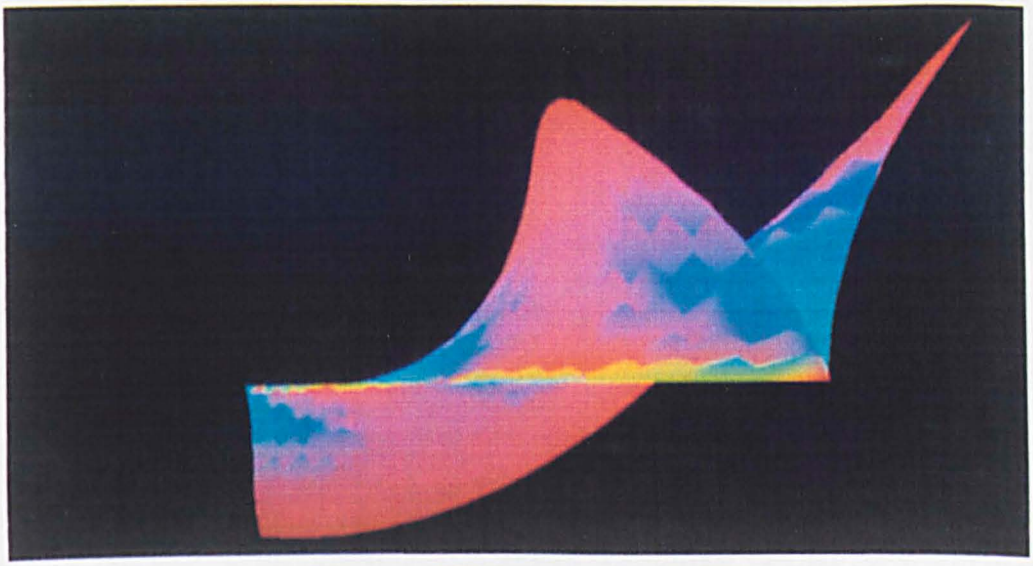
fig. 7.5) The different ways a cuspidal edge can curve.



a) $\kappa = \tau = 0$, g = constant.



b) $g = \tau = 0$, κ = constant.



c) $\kappa = g = 0$, τ = constant.

fig. 7.5) The different ways a cuspidal edge can curve.

Red - negative Gaussian curvature,

Blue - large negative Gaussian curvature.

These pictures here are only the first attempts at drawing focal surfaces as solid coloured surfaces and some of the colouring can be misleading. This is particularly evident in figure 7.5)* where the edge appears yellow and there are sudden changes between red and blue. Figures 7.6)–7.8) have been included to help clarify these picture. The reader is referred to §9.13 where the production of these picture is discussed in detail.

1) The surface $(x, y, z) = (s, t^3, t^2)$ (fig. 7.4). This is the standard model of a cuspidal edge but is very non generic in our study as $\kappa = g = \tau = 0$. At all points of the surface one of the principal curvatures is zero so the Gaussian curvature is zero.

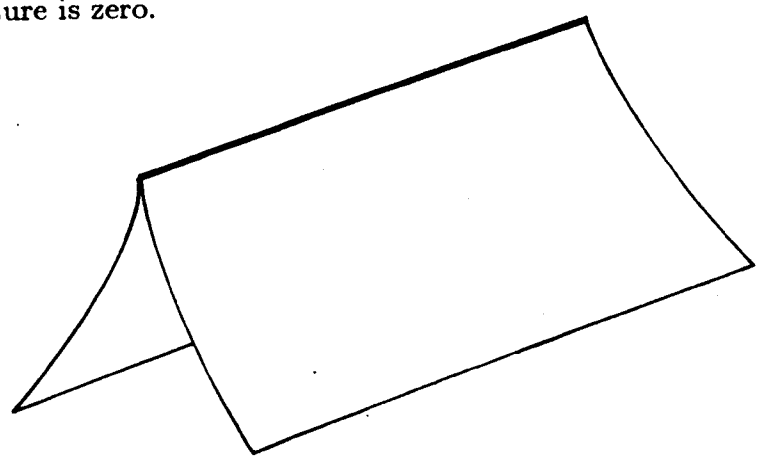


fig. 7.4) A straight cuspidal edge

2) In cylindrical coordinates $f(t, \theta) = (t^2 + r_0)\underline{\hat{r}} + t^3\underline{z}$ defines a cusp edge (fig. 7.5a)*, 7.6). Here we have $\kappa = \tau = 0$. The two families of curves $I_t(\theta) = f(t, \theta)$ and $J_\theta(t) = f(t, \theta)$ intersect at right angles. Furthermore I_t', J_θ' are principal directions and their principal curvatures are

$$\kappa_t = \frac{-3t}{(t^2 + r_0)\sqrt{4 + 9t^2}}, \quad \kappa_\theta = \frac{-6}{t(4 + 9t^2)^{3/2}}$$

The Gaussian curvature is

$$K = \kappa_t \kappa_\theta = \frac{18}{(t^2 + r_0)(4 + 9t^2)^2}$$

which is of constant sign and tends to a *finite* limit as t tends to zero. We note that $\kappa_t \to 0$ as $t \to 0$ and we see that moving along the curve I_0 the normal, $\underline{B}$, does not change direction so we would expect a zero principal curvature.

3) Again in cylindrical coordinates we have a cuspidal edge (fig. 7.5b)*, fig. 7.7)) defined by $f(t,\theta) = (t^3 + r_0)\hat{\underline{r}} + t^2\hat{\underline{z}}$. The two principal curvatures are

$$\frac{2}{(t^3 + r_0)\sqrt{4 + 9t^2}}, \quad \frac{-6}{t(4 + 9t^2)^{3/2}}$$

and their product is

$$K = \frac{-12}{t(t^3 + r_0)(4 + 9t^2)^2}$$

which tends to infinity as t tends to zero and the sign changes as we cross the cuspidal edge. Here $\tau = g = 0$ and κ is non zero.

4) Here we have a straight cuspidal edge (fig 7.5c)*, 7.8)) where $\kappa = g = 0$ but the surface twists around it. First define a basis $(\hat{U}(s), \hat{V}(s), \hat{Z})$ where Z points along the edge and $\hat{U}$, $\hat{V}$ rotate around it, $\hat{U}(s) = \cos(s)\hat{U}(0) + \sin(s)\hat{V}(0)$, $\hat{V}(s) = -\sin(s)\hat{U}(0) + \cos(s)\hat{V}(0)$ (fig. 7.8). The surface is defined by $f(s,t) = s\hat{Z} + t^2\hat{U}(s) + t^3\hat{V}(s)$. As t tends to zero the Gaussian curvature can be shown to tend to a finite negative limit.

5) Now consider a cuspidal edge $\gamma(s)$ with κ, g, τ all non zero constants. We have

$$\begin{pmatrix} \underline{A}' \\ \underline{B}' \\ \underline{C}' \end{pmatrix} = \begin{pmatrix} 0, & \kappa, & g \\ -\kappa, & 0, & \tau \\ -g, & -\tau, & 0 \end{pmatrix} \begin{pmatrix} \underline{A} \\ \underline{B} \\ \underline{C} \end{pmatrix}$$

where primes denote differentiation along the edge. A cuspidal edge is defined by

$$f(s,t) = \gamma(s) + t^3\underline{B} + t^2\underline{C}.$$

If $\underline{e}_1 = (1,0)$, $\underline{e}_2 = (0,1)$ then the first derivatives are

$$df{<}\underline{e}_1{>} = \underline{A} + t^3\underline{B}' + t^2\underline{C}', \qquad df{<}\underline{e}_2{>} = 3t^2\underline{B} + 2t\underline{C},$$

so

$$E = 1 + O(t^2), \qquad F = -t^4\tau, \qquad G = 4t^2 + 9t^4$$

and $EG - F^2 = 4t^2 + O(t^4)$. Now

$$df{<}\underline{e}_1{>} \wedge df{<}\underline{e}_2{>} = 2t\underline{B} - 3t^2\underline{C} + O(t^3)$$

and a unit normal is given by

$$\begin{aligned} \underline{n} &= \frac{2\underline{B} - 3t\underline{C} + O(t^2)}{\sqrt{4 + O(t^2)}} \\ &= \underline{B} - \tfrac{3}{2}t\underline{C} + O(t^2). \end{aligned}$$

Calculating the second derivatives gives

$$d^2f<\underline{e}_1^2> = \underline{A}' + O(t^2), \quad d^2f<\underline{e}_1\underline{e}_2> = 2t\underline{C}' + 3t^2\underline{B}', \quad d^2f<\underline{e}_2^2> = 6t\underline{B} + 2\underline{C}.$$

Hence

$$l = -\kappa + \tfrac{3}{2}gt + O(t^3), \qquad m = 2t\tau + O(t^3), \qquad n = -3t + O(t^2)$$

so the Gaussian curvature is

$$\begin{aligned} K &= \frac{3t\kappa - \frac{9}{2}t^2g - 4\tau^2t^2 + O(t^3)}{4t^2 + O(t^4)} \\ &= \tfrac{3}{4}\kappa t^{-1} - \tfrac{9}{8}g - \tau^2 + O(t). \end{aligned}$$

This shows that if $\kappa \neq 0$ the curvature changes sign through infinity as we cross a cuspidal edge. If $\kappa = 0$ the the curvature will tend to a finite limit and have the same sign on both sides of the cuspidal edge.

From these examples we can see how the numbers κ, g, τ are related to the way in which the cusp edge curves and how κ relates to the Gaussian curvature. The next lemma highlights the connection between $\overline{n}$ and κ for focal surfaces and shows they are both generically non zero.

Lemma 7.10

For a cuspidal edges on a focal surface κ is zero if and only if $\overline{n} = d\underline{P}<q> \cdot \underline{Q}$ is zero or we are at an umbiliç or A_4 point. This will generically happen only at isolated points.

Proof

As we are on a cuspidal edge $df<p> = 0$. The expression for κ can be expanded by writing r as $\alpha p + \beta q$ to get

$$\begin{aligned} \kappa &= d\underline{B}<r> \cdot df<r> \\ &= d\underline{P}<r> \cdot df<r> \\ &= \beta d\underline{P}<r> \cdot df<q> \\ &= \alpha\beta m + \beta^2 n \\ &= \beta^2\ d^2V<q^2>\ \overline{n}. \end{aligned}$$

So κ is zero if and only if $\beta = 0$ or $\overline{n} = 0$ or $d^2V<q^2> = 0$. Now $\beta = 0$ implies that the tangent to the ridge r in the parameter space is in the p direction and from lemma 7.7 we have that this is the condition for an A_4 singularity. If $d^2V<q^2> = 0$ then we have an umbilic point. Umbilics centres and A_4 points are generically isolated points on the ribs, also $l = \overline{n} = 0$ only occurs generically at isolated points.

Chapter 8: Examples of Umbilics

In this chapter we obtain some computer pictures of the pattern of ridges and sub-parabolic lines around umbilics on some surfaces. We also show pictures of focal surfaces which have the ribs and parabolic lines marked on them. The pictures are all generated from generic examples without any symmetry and they are the first time that the (sub-) parabolic lines have been illustrated. Examples of most types of umbilic and some of the transitions are included. Before we look at the examples the classification of umbilics is reviewed in §8.1 and §8.2. Section 8.2 also contains some new results about sub-parabolic lines and in particular their relationship with the lines of curvature and their directions at umbilics.

§8.1 Cubic forms

When studying umbilics there are two cubic forms of interest $d^3V<p^3>$ and $d^3V<pq^2>$. The following section deals with the classification of such forms. The results here have been adapted from [Porteous-3]. We extend these results to classify symmetric cubic forms as well.

Definition 8.1

A **cubic form** is a homogeneous polynomial map $\mathbf{R}^n \to \mathbf{R}$ of degree 3. Likewise **a quadratic form** is a homogeneous polynomial map of degree 2. All forms in this chapter have $\mathbf{R}^2$ as the domain.

Any quadratic form can be written as $Q(x,y) = (\alpha x^2 + 2\beta xy + \gamma y^2)$ where α, β, γ are real numbers. A cubic form can be written as $C(x,y) = (ax^3 + 3bx^2y + 3cxy^2 + dy^3)$, where a, b, c, d are real numbers.

Definition 8.2

A form F has a root direction $u = (x,y)$ if $F(\alpha u) = 0$ for all real values of α.

The number of real root directions determines the type of the form either elliptic, parabolic or hyperbolic.

Proposition 8.3

The quadratic form $Q(x,y) = \alpha x^2 + 2\beta xy + \gamma y^2$ is

hyperbolic: two real roots, if $\beta^2 > \alpha\gamma$,
parabolic: two coincident real root directions, if $\beta^2 = \alpha\gamma$,

elliptic: two complex conjugate root directions (in $\mathbb{C}^2 \supset \mathbb{R}^2$), if $\beta^2 < \alpha\gamma$.

All cubic forms have at least one real root direction and we can have the following types:

elliptic: with three distinct real root directions, if it has a hyperbolic quadratic form as a factor.

parabolic: three real roots two of which are coincident, if it has a parabolic quadratic form as a factor.

cubical: three coincident real root directions, if the form can be written as $C(x,y) = (\lambda x + \mu y)^3$.

hyperbolic: One real root and two complex conjugate roots if it has an elliptic quadratic form as a factor.

We can think of a cubic form C as a third derivative d^3V evaluated on p^3 i.e.

$$C(x,y) = d^3V<p^3> \quad \text{where } p = (x,y).$$

We can also associate the cubic form with a quadratic form

$$H = (ac - b^2)x^2 + (ad - bc)xy + (bd - c^2)y^2$$

called the **hessian**. If u is a root direction of the Hessian then the quadratic form $d^3V<up^2>$ is parabolic. The slightly strange naming of cubic forms comes from the classification of their associated hessians. If the cubic form is elliptic, resp. parabolic, hyperbolic then the hessian is elliptic, resp. parabolic,hyperbolic. The results in this section can be derived through considering the hessian [Porteous–3] pp64–66, although we will not do so here.

Definition 8.4

An **orthogonal form** is one where two of the real root directions are at right angles. A hyperbolic quadratic form is orthogonal if it is of the type $(\lambda x + \mu y)(\mu x - \lambda y) = \lambda\mu x^2 + (\mu^2 - \lambda^2)xy - \lambda\mu y^2$. Elliptic or parabolic cubic forms are orthogonal if they have orthogonal quadratic forms as factors. Hyperbolic cubic forms cannot be orthogonal.

A form $F : \mathbb{R}^2 \to \mathbb{R}$ is **symmetric** if there exists a reflection $R : \mathbb{R}^2, 0 \to \mathbb{R}^2, 0$ such that $F \circ R = F$. This implies that for each root direction u its reflection $R(u)$ is also a root direction. A symmetric cubic form can be factored into symmetric linear form and quadratic forms with the same axis of symmetry. A symmetric linear form, and hence the cubic form, has a root direction pirpendicular to the axis of symmetry.

Definition 8.5 Complex cubic forms

We can identify $\mathbf{R}^2$ with $\mathbf{C}$ by letting $z = (x+iy)$. The quadratic form $Q(x,y) = Ax^2 + 2Bxy + Cy^2$ is equal to the **complex quadratic form**

$$Q^*(z) = \gamma z^2 + 2rz\bar{z} + \bar{\gamma}\bar{z}^2$$

where $\gamma = \frac{1}{4}(A - C - 2Bi) \in \mathbf{C}$ and $r = \frac{1}{4}(A + C) \in \mathbf{R}$. The cubic form

$$C(x,y) = ax^3 + 3bx^2y + 3cxy^2 + dy^3$$

is equal to **the complex cubic form**

$$C^*(z) = \alpha z^3 + 3\bar{\beta}z^2\bar{z} + 3\beta z\bar{z}^2 + \bar{\alpha}\bar{z}^3 \tag{8.1}$$

where $\alpha = \frac{1}{8}(a - 3c + (d - 3b)i)$ and $\beta = \frac{1}{8}(a + c + (d + b)i)$.

Definition 8.6 Equivalence of forms

Two forms F, G are equivalent if there is some linear transformation $T : \mathbf{R}^2, 0 \to \mathbf{R}^2, 0$ consisting of rotations and enlargements only, such that $F = G \circ T$. For complex forms such a transformation can be written as $T(z) = \omega z$, $\omega \in \mathbf{C}$.

If $\alpha \neq 0$ the cubic form (8.1) is equivalent to

$$C_1^*(w) = w^3 + 3\bar{\beta}'w^2\bar{w} + 3\beta' w\bar{w}^2 + \bar{w}^3$$

where $w = \omega z$, $\beta' = \frac{\beta}{\omega\bar{\omega}^2}$ and ω is one of the cube roots of α. We can classify such cubic forms by the position of β' in the plane. If $\alpha = 0$ the form is represented by a point at infinity. We now examine the regions of the β'-plane to see which types of cubic form they correspond to.

Lemma 8.7

The cubic form $C^*(z) = z^3 + 3\bar{\beta}z^2\bar{z} + 3\beta z\bar{z}^2 + \bar{z}^3$ is

i) cubical if $\beta = 1, e^{\frac{2\pi i}{3}}, e^{\frac{-2\pi i}{3}}$;

ii) parabolic if $\beta \in \Gamma = \{\frac{1}{3}(2e^{i\theta} + e^{-2i\theta})\}$ and the form is not cubical;

iii) elliptic if β lies inside Γ;

iv) hyperbolic if β lies outside Γ;

v) orthogonal if $|\beta| = \frac{1}{3}$;

vi) symmetric if $\beta = re^{\frac{2n\pi i}{3}}$, $r \in \mathbf{R}$, $n \in \mathbf{Z}$.

See figure 8.1.

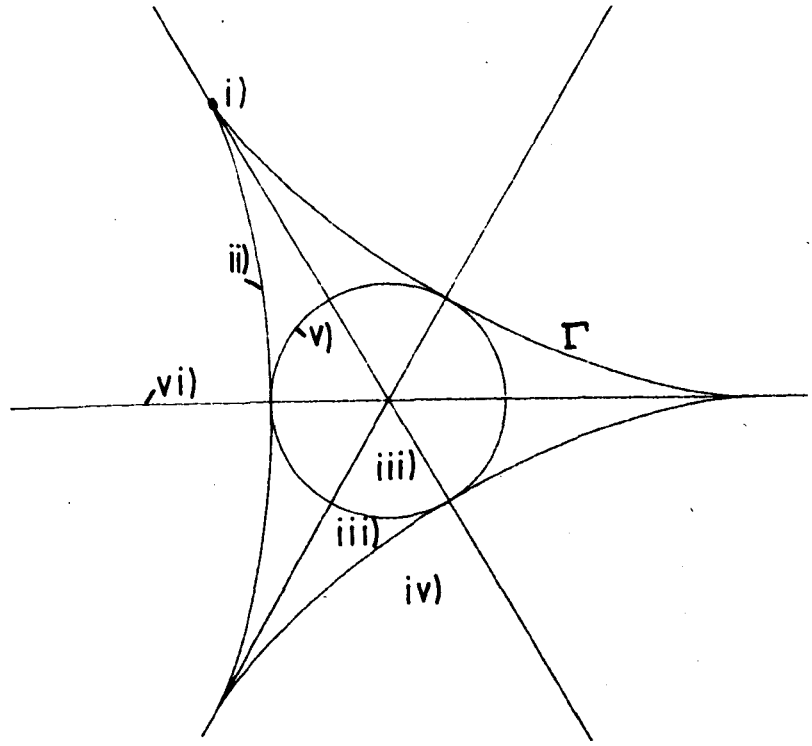


fig. 8.1) Classification of cubic forms in lemma 8.7

Proof

Adapted from [Porteous-3] pp65–66. Let $Q^*(z) = e^{i\theta}z^2 + 2rz\bar{z} + e^{-i\theta}\bar{z}^2$ be a complex quadratic form. Any quadratic form is equal to a real multiple of a quadratic form of this type. This form is hyperbolic if $|r|<1$, parabolic if $r = 1$ and elliptic if $|r|>1$. The cubic form C^* must have a quadratic factor of this type, the other factor being linear: $L^*(z) = (e^{i\phi}z + e^{-i\phi}\bar{z})$. Now

$$Q^*(z)L^*(z) = e^{i(\theta+\phi)}z^3 + (2re^{i\phi} + e^{i(\theta-\phi)})z\bar{z}^2 + (2re^{-i\phi} + e^{i(\phi-\theta)})z\bar{z}^2 + e^{-i(\theta+\phi)}\bar{z}^3.$$

Equating the coefficients of z^3 with $C^*(z)$ gives $\phi = -\theta$ and equating coefficients of $z\bar{z}^2$ shows that $3\beta = e^{-2i\theta} + 2re^{i\theta}$.

For $r = 1$ this gives the curve Γ. Points on this curve will correspond to cubical or parabolic forms. Elliptic forms will have $|r|<1$ and correspond to points inside Γ. Hyperbolic forms have $|r|>1$ and correspond to points outside the curve. In particular orthogonal quadratic forms are of the form $e^{i\theta}z^2 + e^{-i\theta}\bar{z}^2$ and have $r = 0$. These will give points on the circle $\beta = \frac{1}{3}e^{2i\theta}$.

A parabolic cubic form will be cubical when its quadratic factor is the square of the linear factor i.e.

$$e^{2i\theta}z^2 + 2z\bar{z} + e^{-2i\theta}\bar{z}^2 = (e^{-i\theta}z + e^{i\theta}\bar{z})^2,$$

hence $e^{i\theta} = e^{-2i\theta}$ so $\theta = \frac{2n\pi}{3}$ for some integer n. This gives the three points at the cusps of Γ.

A reflection in a line through the origin at an angle ψ to the x-axis is defined by

$$R_\psi : z \to e^{2\psi i}\bar{z}.$$

For symmetric forms $C^* \circ R_\psi = C^*$ for some ψ. Now $C^* \circ R_\psi$ is

$$e^{6i\psi}\bar{z}^3 + 3\bar{\beta}e^{2i\psi}\bar{z}^2 z + 3\beta e^{-2i\psi}z^2\bar{z} + e^{-6i\psi}z^3.$$

Equating coefficients with C^* gives $\psi = \frac{n\pi}{3}$, $\beta = \rho e^{i\psi + m\pi}$, $m, n \in \mathbf{Z}$, $\rho \in \mathbf{R}$. This gives three straight lines through the origin.

§8.2 Umbilic classification

We now examine the link between umbilics and the cubic forms which describe their properties. There are two cubic forms which describe the pattern of ridges and sub-parabolic lines around an umbilic, $C_1(x, y) = d^3V{<}p^3{>}$ and C_2 the Jacobian of C_1 and the first fundamental form of the surface. The ridges are related to C_1 and C_2 describes the sub-parabolic lines and line of curvature. Let $s : \mathbf{R}^2 \to \mathbf{R}^3$ be a parametrization of a surface with $s(0)$ an umbilic point and let $V : \mathbf{R}^2 \times \mathbf{R}^3 \to \mathbf{R}$ be the associated distance squared map. Without loss of generality we assume that the first fundamental form at the umbilic, $I_0{<}u, v{>}$, is just the dot product on $\mathbf{R}^2$, i.e. $I_0{<}u, v{>} = u \cdot v$. For any direction r we write $r^\perp$ for a vector perpendicular to r i.e. $I{<}rr^\perp{>} = 0$. In particular if p is a principal direction then $p^\perp$ is also a principal direction and we write q for $p^\perp$.

If $\gamma : \mathbf{R} \to \mathbf{R}^2$ is smooth curve such that $\gamma(0) = 0$ then we can find a pair of principal direction p_t, q_t at each point t of the curve except 0. We now wish to find out if there is a well defined limit for p_t, q_t as $t \to 0$. (We recall that every direction is principal at an umbilic point).

Lemma 8.8

The principal directions p_t, q_t have limits p_0, q_0 as $t \to 0$ which satisfy

$$d^3V{<}p_0q_0r{>} = 0$$

where $r = \frac{d\gamma}{dt}(0)$ is the tangent to γ at $\gamma(0)$.

Proof

We are only interested in the direction of p_0 and not its length. The length can be eliminated by writing $p_t = p_{xt}\underline{e}_1 + p_{yt}\underline{e}_2$ and finding an expression for p_{xt}/p_{yt}. At all non umbilical points $d^2V<p_t*> = 0$. Expressing in coordinates gives

$$p_{xt}d^2V<\underline{e}_1*> + p_{yt}d^2V<\underline{e}_2*> = 0$$

and

$$\frac{p_{xt}}{p_{yt}} = \frac{-d^2V<\underline{e}_2*>}{d^2V<\underline{e}_1*>}.$$

Both numerator and denominator are zero at $t = 0$ so we use L'Hopitals rule to find the limit. This gives

$$\frac{p_{x0}}{p_{y0}} = \frac{-d^3V<\underline{e}_2 r*> + d^2s<\underline{e}_2*> \cdot df<r>}{d^3V<\underline{e}_1 r*> - d^2s<\underline{e}_1*> \cdot df<r>} \qquad (8.2).$$

This equation holds for any vector substituted for $*$. Provided $d^3V<r**> \neq d^2s<**> \cdot df<r>$ we can find the limit of p_{xt}/p_{yt}. Rearranging equation 8.2 gives

$$d^3V<p_0 r*> = d^2s<p_0*> \cdot df<r>.$$

In particular inserting q_0 proves the theorem.

Lemma 8.9

If the limiting principal direction p_0 satisfies $d^3V<p_0 q_0^2> \neq 0$ then $d\underline{P}<q> \cdot \underline{Q}$ tends to infinity as we approach the umbilic along a curve γ.

Proof

Differentiating $d^2V<pq> = 0$ in the q-direction gives

$$d^3V<pq^2> + d^2V<dp<q>q> - d^2s<pq> \cdot df<q> = 0.$$

This equation holds at all points except the umbilic. At each point we write $dp<q>$ as $\alpha<q>p + \beta<q>q$ where α, β are linear maps $\mathbf{R}^2 \to \mathbf{R}$. From the above equation we find that

$$\beta<q> = \frac{d^2s<pq> \cdot df<q> - d^3V<pq^2>}{d^2V<q^2>}.$$

We now restrict our attention to the curve γ and parametrize everything by t to have maps $I(t)$, $II(t)$, $ds(t)$, $df(t)$, $d^2V(t)$, $d^3V(t)$ and vector fields p_t, q_t, $\underline{P}_t$,

$\underline{Q_t}$, $\underline{n_t}$. We can simplify $df(t)<q_t>$ by expressing it in terms of the $\underline{P_t}$, $\underline{Q_t}$, $\underline{n_t}$ basis. We have

$$df(t)<q_t> = a\underline{Q_t} + b\underline{n_t}$$

for some a, $b \in \mathbf{R}$. This follows as $\underline{P_t}$ is normal to the focal surface (lemma 7.3). From equation 7.6 we have $a = df(t)<q_t> \cdot \underline{Q_t} = d^2V(t)<q_t^2>$. Now calculating $\beta<q_t>$ we have

$$\beta<q_t> = \frac{d^2V(t)<q_t^2>\, d^2s(t)<p_tq_t> \cdot \underline{Q_t} - d^3V(t)<p_tq_t^2>}{d^2V(t)<q_t^2>}.$$

Taking limits as $t \to 0$ we have $d^2V(t)<**> \to 0$ and

$$\lim_{t\to 0} \beta_t<q_t> = d^2s(0)<p_0, q_0> \cdot \underline{Q_0} - \lim_{t\to 0} \frac{d^3V(t)<p_tq_t^2>}{d^2V(t)<q_t^2>}.$$

Now if $d^3V(0)<p_0q_0^2> \neq 0$ this will have an infinite limit.

Differentiating $\underline{P} = ds<p>$ in the q direction gives

$$\begin{aligned} d\underline{P}<q> \cdot \underline{Q} &= d^2s<pq> \cdot \underline{Q} + ds<dp<q>> \cdot \underline{Q} \\ &= d^2s<pq> \cdot \underline{Q} + \beta<q>. \end{aligned}$$

Now parametrizeing by t and examining the limit as $t \to 0$ proves the lemma.

Corollary 8.10

If the limit of $d\underline{P}<q> \cdot \underline{Q}$ is finite as we approach an umbilic along a curve γ then the limiting principal directions satisfy $d^3V<p_0q_0^2> = 0$. In particular this holds for sub-parabolic lines where $d\underline{P}<q> \cdot \underline{Q} = 0$.

One interesting family of curves are the **lines of curvature**. These curves are defined by the differential equation $\frac{d\gamma}{dt} = p_t$. There are two sets of lines of curvature, p-lines and q-lines. At every non-umbilic point one line from the first set crosses a line from the second set at right angles. Around an umbilic there are three generic patterns of lines of curvature "Star", "Monstar", and "Lemon" first discovered in [Darboux]. The name are from [Hannay] (fig 8.2).

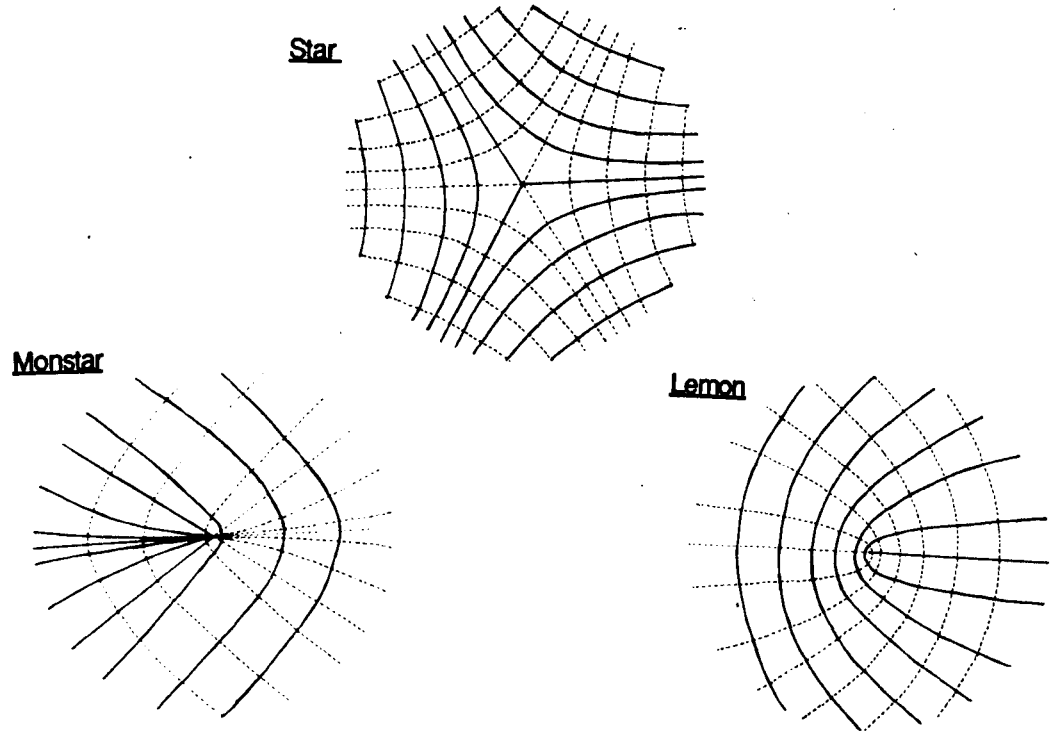


fig. 8.2) Pattern of lines of curvature around an umbilic (standard pictures)

The **Geodesic curvature** g of a curve C, on a surface at a point is the curvature of the curve when projected on to the tangent plane. It can be calculated from $g = \underline{T}' \cdot \underline{V}$, where $\underline{T}$ is the tangent to the curve, $\underline{V} = \underline{T} \wedge \underline{n}$ is at right angles to $\underline{T}$ in the tangent plane and $\underline{T}'$ is the derivative of $\underline{T}$ with respect to arc length (fig 8.3a).

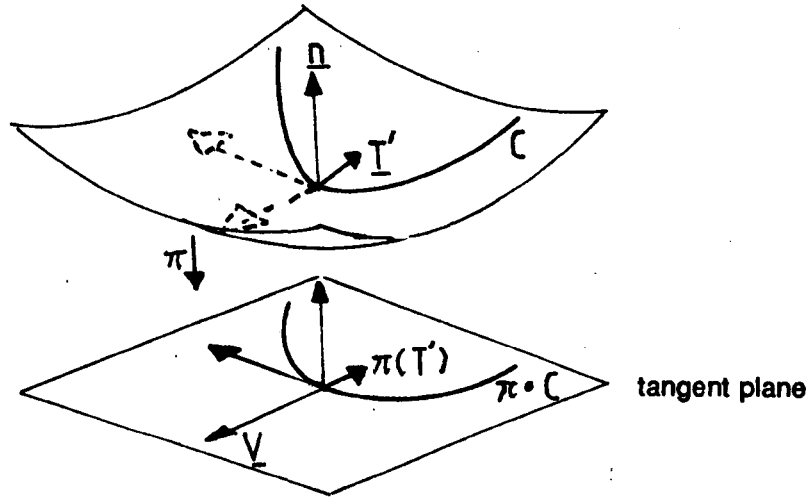


fig. 8.3a) Projection of curve onto tangent plane

Theorem 8.11

The geodesic curvature of a q-line of curvature at a point x is zero if and only if x is on a sub-parabolic line of the p-sheet of the focal surface.

Proof

For a q-line of curvature the geodesic curvature is given by

$$\begin{aligned} g &= d\underline{Q}<q> \cdot \underline{P} \\ &= -d\underline{P}<q> \cdot \underline{Q} \end{aligned}$$

so is zero precisely when we are on a sub parabolic line.

If we approach an umbilic along a curve C we find that unless $d^3V<p_0q_0^2> = 0$ the geodesic curvature of each successive q-line of curvature we cross will tend to infinity (fig. 8.3b). We note that in the "Star" and "Monstar" patterns there are three **Exceptional directions** along which lines of curvature with finite curvature pass through the umbilic. In the "Lemon" pattern there is only one such directions. The number of these direction is controlled by the cubic form $d^3V<p_0q_0^2>$. We note in figure 8.2 none of the lines of curvature have inflections. We will see later that generically there is at least one sub-parabolic line going through the umbilic so we would expect the lines of curvature to have inflections as in figure 8.3b).

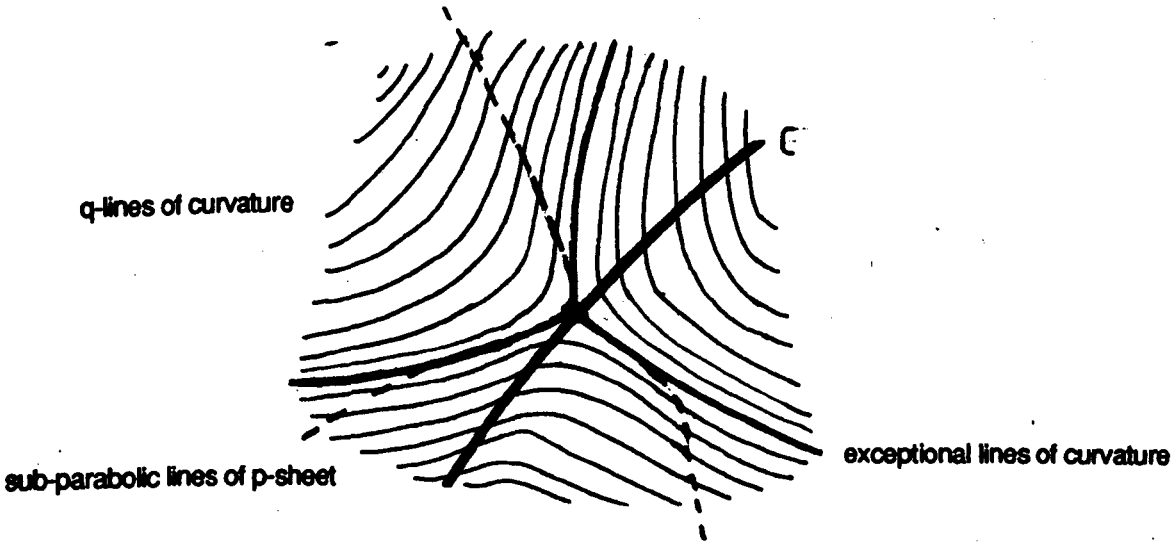


fig. 8.3b) Sub-parabolic lines and exceptional lines of curvature around an umbilic.

Proposition 8.12

At an umbilic point we have the following results.

1) Provided the cubic $d^3V<u^3>$ is not parabolic the ridges pass smoothly through the umbilic and their tangents, r, satisfy $d^3V<pqr> = 0$ where p is a solution of $d^3V<p^3> = 0$. Also the ribs pass smoothly through the umbilic centre

and they are tangent to the normal line to the original surface S. If $d^3V<u^3>$ is parabolic then one of the ridges and the corresponding rib are both singular.

2) If the cubic $d^3V<u(u^\perp)^2>$ is neither orthogonal or parabolic then the sub-parabolic lines pass smoothly through the umbilic with tangent r satisfying $d^3V<pqr> = 0$ where p is a solution of $d^3V<pq^2> = 0$. This implies $r = q$ and $d^3V<r^2r^\perp> = 0$. On the focal surface the parabolic lines pass smoothly through the umbilic centre and are tangent to the normal line to the original surface. If $d^3V<u(u^\perp)^2>$ is parabolic then one of the parabolic lines and the corresponding sub-parabolic line are both singular.

3) Provided $d^3V<u^2u^\perp>$ is not orthogonal or parabolic then the exceptional p-lines of curvature pass smoothly through the umbilic with tangent $r = p$ satisfying $d^3V<r^2r^\perp> = 0$. Hence they are tangent to the sub-parabolic lines.

Proof

Part 1) of this proof is inspired by Ian Porteous which has been adapted for parts 2), 3).Consider the map

$$M : \mathbf{R}^2 \times S^1 \times \mathbf{R}^3 \to [\mathbf{R}^2, \mathbf{R}] \times [\mathbf{R}^2, \mathbf{R}] \cong \mathbf{R}^2 \times \mathbf{R}^2$$

$$M : (x,p,f) \to (dV<*>, d^2V<p*>) = (dV<p>, dV<q>, d^2V<p^2>, d^2V<pq>).$$

Let $\pi_x : (x,p,f) \to x$, $\pi_p : (x,p,f) \to p$, and $\pi_f : (x,p,f) \to f$ be the three canonical projection maps. If $(x,p,f) \in M^{-1}(0)$ then $\pi_x(x,p,f)$ is a point in parameter space with associated principal direction $\pi_p(x,p,f)$ and focal point $\pi_f(x,p,f)$. We can write a tangent vector to $\mathbf{R}^2 \times S^1 \times \mathbf{R}^3$ as $(r_1p + r_2q,\ \zeta q,\ f_1\underline{P} + f_2\underline{Q} + f_3\underline{n})$ where r_1, r_2, ζ, f_1, f_2 and f_3 are real numbers. Note p is of unit length so its derivative will be in the q direction.

The derivative of M is a linear map between the tangent spaces and

$$dM<r_1, r_2, \zeta, f_1, f_2, f_3>$$

$$= \begin{pmatrix} d^2V<p^2>, & d^2V<pq>, & 0, & \underline{\lambda}\cdot\underline{P}, & \underline{\lambda}\cdot\underline{Q}, & \underline{\lambda}\cdot\underline{n} \\ d^2V<pq>, & d^2V<q^2>, & 0, & \underline{\mu}\cdot\underline{P}, & \underline{\mu}\cdot\underline{Q}, & \underline{\mu}\cdot\underline{n} \\ d^3V<p^3>, & d^3V<p^2q>, & d^2V<pq>, & \underline{\nu}\cdot\underline{P}, & \underline{\nu}\cdot\underline{Q}, & \underline{\nu}\cdot\underline{n} \\ d^3V<p^2q>, & d^3V<pq^2>, & d^2V<q^2>, & \underline{\xi}\cdot\underline{P}, & \underline{\xi}\cdot\underline{Q}, & \underline{\xi}\cdot\underline{n} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ \zeta \\ f_1 \\ f_2 \\ f_3 \end{pmatrix}$$

$$= \begin{pmatrix} 0, & 0, & 0, & \underline{\lambda}\cdot\underline{P}, & 0, & 0 \\ 0, & d^2V<q^2>, & 0, & 0, & \underline{\mu}\cdot\underline{Q}, & 0 \\ d^3V<p^3>, & d^3V<p^2q>, & 0, & \underline{\nu}\cdot\underline{P}, & \underline{\nu}\cdot\underline{Q}, & \underline{\nu}\cdot\underline{n} \\ d^3V<p^2q>, & d^3V<pq^2>, & d^2V<q^2>, & \underline{\xi}\cdot\underline{P}, & \underline{\xi}\cdot\underline{Q}, & 0 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ \zeta \\ f_1 \\ f_2 \\ f_3 \end{pmatrix}.$$

Here $\underline{\lambda} = -ds<p>$, $\underline{\mu} = -ds<q>$, $\underline{\nu} = -d^2s<p^2>$ and $\underline{\xi} = -d^2s<pq>$. Tangent vectors to M lie in the kernel of this map. At all non umbilic points the kernel has dimension two.

At an umbilic $d^2V<q^2> = 0$ and we can assume that we have a Monge parametrization with $I<uv> = u \cdot v$. Hence $f_1 = f_2 = 0$ and the other components lie in the kernel of

$$\begin{pmatrix} d^3V<p^3>, & d^3V<p^2q>, & 0, & -\kappa \\ d^3V<p^2q>, & d^3V<pq^2>, & 0, & 0 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ \zeta \\ f_3 \end{pmatrix} = 0,$$

where κ is the principal curvature at the umbilic. This has a two dimensional kernel provided $d^3V<pqu> \neq 0$ for some vector u. Hence the set $M^{-1}(0)$ is generically a smooth manifold of dimension 2 even at an umbilic. We can think of this map as a desingularisation or blow up of the umbilic.

1) A ridge and its associated rib can be considered as a curve γ_R in $M^{-1}(0)$ for which $d^3V<p^3> = 0$; the ridge is given by $\pi_x(\gamma_R)$ and the rib by $\pi_f(\gamma_R)$. Differentiating $d^3V<p^3> = 0$ gives

$$(\, d^4V<p^4>, \quad d^4V<p^3q>, \quad 3d^3V<p^2q>, \quad \underline{\rho} \cdot \underline{P}, \quad \underline{\rho} \cdot \underline{Q}, \quad \underline{\rho} \cdot \underline{n}\,) \begin{pmatrix} r_1 \\ r_2 \\ \zeta \\ f_1 \\ f_2 \\ f_3 \end{pmatrix} = 0,$$

where $\underline{\rho} = d^3s<p^3>$. Now at an umbilic we have the following conditions

$$d^3V<p^2r> - \kappa f_3 = 0 \tag{8.3}$$

$$d^3V<pqr> = 0 \tag{8.4}$$

$$d^4V<p^3r> + 3d^3V<p^2q>\zeta = 0 \tag{8.5}$$

As we have assumed that we have a Monge form parametrization we have $d^3s<p^3> \cdot \underline{n} = \kappa\, d^3V<p^3> = 0$.

If $d^3V<p^2q> \neq 0$ then we can find a non zero r satisfying $d^3V<pqr> = 0$ and also non zero numbers ζ, f_3 which satisfy equations 8.3 and 8.5. Examining the projections shows that ribs and ridges are non singular.

If $d^3V<p^2q> = 0$ then $d^4V<p^3r> = 0$. Generically this does not have a simultaneous solution with $d^3V<pqr> = 0$ so $r = 0$ and from equation (8.3) $f_3 = 0$.

Examining the projections $\pi_x(\gamma_R)$ and $\pi_f(\gamma_R)$ shows that here both rib and ridge are singular.

If we let $u = xp + yq$ then the cubic $d^3V<u^3>$ is given by

$$d^3V<p^3>x^3 + 3d^3V<p^2q>x^2y + 3d^3V<pq^2>xy^2 + d^3V<q^3>y^3.$$

For ribs and ridges $d^3V<p^3> = 0$ and if $d^3V<p^2q> = 0$ the cubic reduces to

$$y^2(3d^3V<pq^2>x + d^3V<q^3>y)$$

which is clearly parabolic.

2) Now consider the curve $\gamma_P \subset M^{-1}(0)$ defined by $d^3V<pq^2> = 0$. Differentiating this equation gives

$$d^4V<pq^2r> + (d^3V<q^3> - d^3V<p^2q>)\zeta = 0. \tag{8.6}$$

As above the projections $\pi_x(\gamma_P)$, $\pi_f(\gamma_P)$ will be singular if

$$d^3V<q^3> - 2d^3V<p^2q> = 0. \tag{8.7}$$

Otherwise the tangent vectors are given by equation (8.6) together with $d^3V<pqr> = 0$, $f_3 = \frac{1}{\kappa}d^3V<p^2r>$ and $d^3V<pq^2> = 0$. Provided $d^3V<pq*> \neq 0$ it follows that $r = q$.

The cubic $d^3V<u(u^\perp)^2>$ is

$$\begin{aligned} d^3V<pq^2>x^3 + (d^3V<q^3> - 2d^3V<p^2q>)x^2y \\ + (d^3V<p^3> - 2d^3V<pq^2>)xy^2 + d^3V<p^2q>y^3. \end{aligned}$$

If $d^3V<pq^2> = 0$ and (8.7) holds this equation reduces to the parabolic cubic form

$$y^2(d^3V<p^3>x + d^3V<p^2q>y).$$

If $d^3V<pq*> = 0$ then we can not determine r from these equations; this implies $d^3V<pq^2> = d^3V<p^2q> = 0$ and that the cubic is orthogonal.

Sub-parabolic lines are defined by

$$d\underline{P}<q> \cdot \underline{Q} = d^3V<pq^2> - 2d^2V<q^2>\, d^2s<pq> \cdot \underline{Q} = 0.$$

Both $d^2s<pq> \cdot \underline{Q}$ and $d^2V<q^2>$ tend to zero as we approach the umbilic so the sub-parabolic lines and the parabolic lines on the focal surface will have the same tangents at the umbilic as $\pi_x(\gamma_P)$ and $\pi_f(\gamma_P)$ respectively.

3) For a line of curvature $r = \alpha p$ so

$$\alpha d^3V<p^3> - \kappa f_3 = 0$$
$$\alpha d^3V<p^2q> = 0.$$

These only have a non zero solution when $d^3V<p^2q> = 0$, which defines the tangent direction of the exceptional lines of curvature.

We now have two cubic forms which describe the possible limiting principal directions at umbilics we now wish to see how these two forms are related to the tri-cusp figure 8.3. For the ridges we use the form $C_1(x,y) = d^3V<(x,y)^3>$ and the principal directions are the root directions of this form. Converting to complex cubic forms and finding the number β classifies the form. For sub-parabolic lines we need to find a relationship between this form and $d^3V<(x,y)(-y,x)^2>$. Let $C_1^*(z) = d^3V<z^3>$, $C_2^*(z) = d^3V<z^2z^\perp>$ and $C_3^*(z) = d^3V<z(z^\perp)^2>$, where we think of $z = x + iy \in \mathbf{C} \cong \mathbf{R}^2$ as tangent vector (x,y) in $\mathbf{R}^2$.

Proposition 8.13

If

$$C_1^*(z) = z^3 + 3\bar{\beta}z^2\bar{z} + 3\beta z\bar{z}^2 + \bar{z}^3$$

then

$$C_2^*(z) = (iz)^3 - \bar{\beta}(iz)^2(i\bar{z}) - \beta(iz)(i\bar{z})^2 + (i\bar{z})^3$$

and

$$C_3^*(z) = z^3 - \bar{\beta}z^2\bar{z} - \beta z\bar{z}^2 + \bar{z}^3$$

Proof

From [Porteous-3] p70. The Jacobian of $C_1^*(z)$ and the first fundamental form, $I<z^2>$, is

$$\begin{vmatrix} \frac{\partial d^3V<z^3>}{\partial x} & \frac{\partial d^3V<z^3>}{\partial y} \\ \frac{\partial I<z^2>}{\partial x} & \frac{\partial I<z^2>}{\partial y} \end{vmatrix} = 6 \begin{vmatrix} d^3V<z^2\underline{e}_1> & d^3<z^2\underline{e}_2> \\ I<z\underline{e}_1> & I<z\underline{e}_2> \end{vmatrix}$$
$$= 6\, d^3V<z^2\,(y\underline{e}_1 - x\underline{e}_2)>$$
$$= 6\, d^3V<z^2z^\perp>$$
$$= 6C_2^*(z).$$

Considering polar coordinates $z = re^{i\theta}$, $x = r\cos(\theta)$, $y = r\sin(\theta)$ the cubic $C_1^*(z)$ can be written as

$$C_1^+(r,\theta) = C_1(r\cos(\theta), r\sin\theta).$$

Calculating the Jacobian

$$6\ C_2^*(z) = \begin{vmatrix} \frac{\partial dC_1(z)}{\partial x} & \frac{\partial dC_1(z)}{\partial y} \\ \frac{\partial I<z^2>}{\partial x} & \frac{\partial I<z^2>}{\partial y} \end{vmatrix} = 2y\frac{\partial dC_1}{\partial x} - 2x\frac{\partial dC_1}{\partial y}$$

$$= -2\frac{\partial C_1^+}{\partial \theta}$$

$$= -2r^3(3ie^{3i\theta} + 3i\bar{\beta}e^{i\theta} - 3i\beta e^{-i\theta} - 3ie^{-3i\theta})$$

$$= -6i(z^3 + \bar{\beta}z^2\bar{z} - \beta z\bar{z}^2 - \bar{z}^3)$$

$$= 6((iz)^3 - \bar{\beta}(iz)^2(i\bar{z}) - \beta(iz)(i\bar{z})^2 + (i\bar{z})^3).$$

If we let $w = iz$ we see the above expression is the equation for $C_3^*(w)$.

We now have a relation between the two cubic forms and can describe their types.

Theorem 8.14

Let U be an umbilic point and let $C_1^*(z) = z^3 + 3\bar{\beta}z^2\bar{z} + 3\beta z\bar{z}^2 + \bar{z}^3$ be a complex cubic form equivalent to $d^3V<z^3>$. The parts of the β-plane illustrated in figure 8.4) correspond to the following types of umbilic.

i) elliptic umbilic, star pattern, 3 ridges, 3 sub-parabolic lines;

ii) as i) but two of the ridges are tangent;

iii) as i) but different arrangement of ridges;

iv) parabolic umbilic, 2 ridges (one is singular), 3 sub-parabolic lines;

v) hyperbolic umbilic, star pattern, 1 ridge, 3 sub-parabolic lines;

vi) circle corresponds to the birth of umbilics 1 ridge, 5 sub-parabolic lines;

vii) hyperbolic umbilic, monstar pattern, 1 ridge, 3 sub-parabolic lines;

viii) hyperbolic umbilic, 1 ridge, 2 sub-parabolic lines (one is singular);

ix) hyperbolic umbilic, lemon pattern, 1 ridge, 1 sub-parabolic line;

x) symmetrical umbilics, one of the ridges and one of the sub-parabolic lines are tangent to the axis of symmetry.

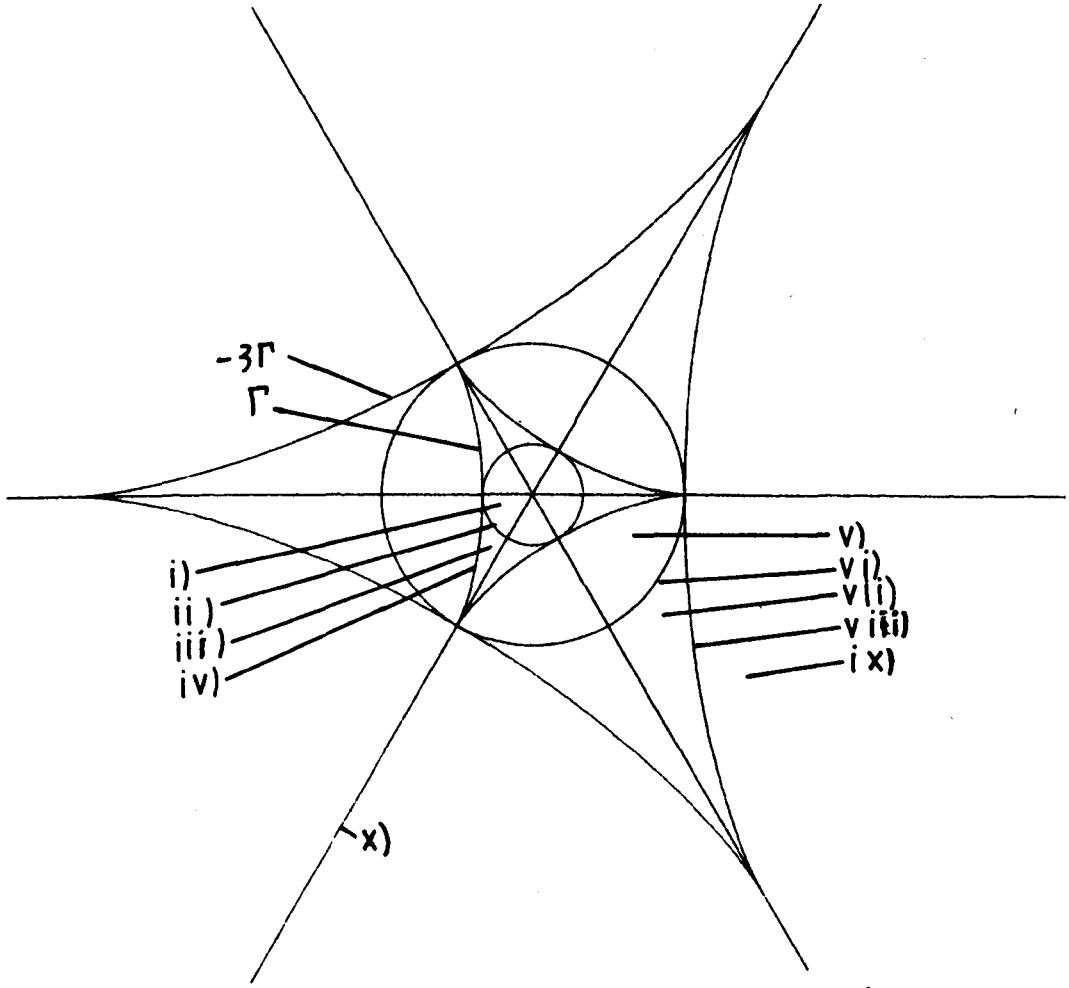


fig. 8.4) Classification of umbilics by position in β-plane

Proof

The arrangement sub-parabolic lines was first discussed in [Bruce–Wilkinson] where they obtain the same classification through considering infinitesimal reflectional symmetry. They also classifies the ridges. Here we follow [Porteous–3] by finding classification in terms of the distance squared function.

The curve $\Gamma = \{\frac{1}{3}(2e^{i\theta} + e^{-2i\theta})\}$ gives the cubic where $C_1^*(z) = d^3V{<}z^3{>}$ is parabolic and -3Γ gives cubics where $C_3(z) = d^3V{<}z(z^{\perp})^2{>}$ is parabolic. When C_1 is orthogonal two the of root direction p_1, p_2 are orthogonal. Now if r_1, r_2 are the tangents to the corresponding ridges we have

$$\begin{aligned} d^3V{<}p_1q_1r_1{>} &= 0, \\ d^3V{<}p_1q_1r_2{>} &= -d^3V{<}q_2p_2r{>} \\ &= 0, \end{aligned}$$

so the two ridges are tangent.

If C_1 is symmetric then C_3 is also symmetric with the same axis of symmetry. This means that p_1, a vector orthogonal to the axis of symmetry, is a root direction for both forms and q_1 is the tangent to both the ridge and the sub-parabolic line.

An interesting proof of the shape of the patterns of lines of curvature has been recently published [Bruce and Fidal]. See also [Sotomayer and Guttierrez].

§8.3 Examples of Umbilics

We now look in closer detail at the various types of umbilics and the transitions which can occur. The program described in Chapter 9 together with the calculations developed in chapter 7 have been used to produce all the following examples. These are divided into two different types.

a) The Projection of the pattern of ridges and sub parabolic lines, which lie on the surface, on to the tangent plane. This is equivalent to the pattern in parameter space.

b) Pictures of focal surface with ribs (cuspidal edges) and parabolic lines drawn on it.

First we look at the pattern of ridges and sub-parabolic lines on the surface. In these examples we will divide the ridges into two different colours according to which sheet of the focal surface the associated ribs lie on. Here black denotes ridges corresponding to ribs on the sheet of the focal surface given by the higher of the two principal curvatures (sheet 1) and red corresponds to ribs on sheet 2. Likewise the sub-parabolic lines are given two colours: blue for the lines corresponding to parabolic lines on sheet 1 and green for sheet 2. At a generic umbilic each ridge or sub-parabolic line will pass smoothly though the umbilic and they will change colour as they do so. Sometimes we will just consider the **half ridges** or **half sub-parabolic lines** which are of just one colour and end at the umbilic.

For many of the transitions involving the ridges and all those involving sub-parabolic lines this is the first time that such pictures have been produced. Full mathematical descriptions and proofs have yet to be found for some of these transitions, the pictures here may form the first step towards discovering such proofs.

Before we study the examples we need to find a method of defining the desired umbilic. If the surface is parametrized in Monge form,

$$s(x,y) = (x, y, \tfrac{1}{2}(Ax^2 + 2Bxy + Cy^2) + \tfrac{1}{6}(ax^3 + 3bx^2y + 3cxy^2 + dy^3) + O(4)),$$

then the derivatives of s at the origin are given by

$$ds<\underline{e}_1> = (1,0,0), \qquad ds<\underline{e}_2> = (0,1,0),$$
$$d^2s<\underline{e}_1^2> = (0,0,A), \quad d^2s<\underline{e}_1\,\underline{e}_2> = (0,0,B), \quad d^2s<\underline{e}_2^2> = (0,0,C),$$
$$d^3s<\underline{e}_1^3> = (0,0,a), \quad d^3s<\underline{e}_1^2\,\underline{e}_2> = (0,0,b), \quad d^3s<\underline{e}_1\,\underline{e}_2^2> = (0,0,c),$$
$$d^3s<\underline{e}_2^3> = (0,0,d).$$

Now using the expression for the derivatives of the distance squared function, V, from §7.1 and measuring the distanee from a point (X,Y,Z) gives

$$V(0,0,X,Y,Z) = \tfrac{1}{2}(X^2+Y^2+Z^2),$$
$$dV<\underline{e}_1> = -X, \qquad dV<\underline{e}_2> = -Y,$$
$$d^2V<\underline{e}_1^2> = 1-AZ, \quad d^2V<\underline{e}_1\,\underline{e}_2> = -BZ, \quad d^2V<\underline{e}_2^2> = 1-CZ.$$

For (X,Y,Z) to be a point on the focal surface corresponding to the the origin of the surface s we require that $X = Y = 0$ and for it to be an umbilic centre the quadratic coefficients must be $B = 0$, $A = C = 1/Z$. At such an umbilic the third derivatives are given by

$$d^3V<\underline{e}_1^3> = aZ, \quad d^3V<\underline{e}_1^2\,\underline{e}_2> = bZ, \quad d^3s<\underline{e}_1\,\underline{e}_2^2> = cZ, \quad d^3s<\underline{e}_2^3> = dZ.$$

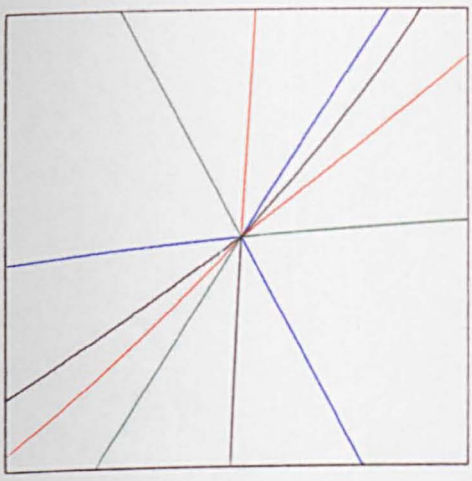
So we see that the coefficients of the cubic $d^3V<(x,y)^3>$ are real multiples of the cubic terms in the Monge parametrization. It is now a simple step to convert to a complex cubic form and plot the point corresponding to the umbilic in the β-plane. Alternatively if we want to specify an umbilic of a given type, we find the position $\beta = s+it$ in the plane. By choosing $\alpha = 1$ in equation 8.1 we have

$$a = 2+6s, \qquad b = 2t, \qquad c = 2s-2, \qquad d = 6t.$$

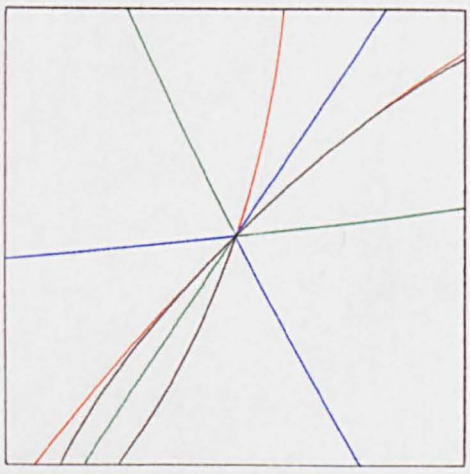
Using these as the coefficients of the cubic terms in the Monge parametrization gives the required umbilic.

§8.4 Orthogonal transition

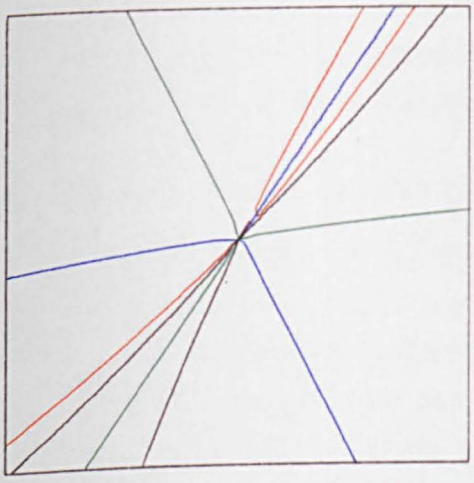
Figure 8.5)* illustrates what happens as we cross the circle $\beta = \frac{1}{3}$. The sub-parabolic lines are unaffected by this transition but we see that two of the ridges become tangent. If we sketch the pattern of ridges we have the following transition (fig. 8.6). Inside the circle we have an alternating pattern of a half ridge of one



1: Inside circle



2: On circle

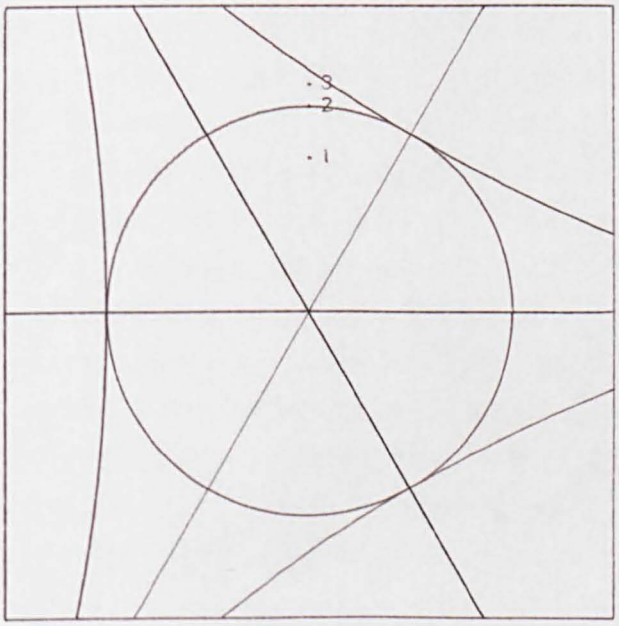


3: Outside circle

Key

- Ridges For sheet 1
- Ridges For sheet 2
- Sub-parabolic lines For sheet 1
- Sub-parabolic lines For sheet 2

Ridges + sub-parabolic lines on surface



Position in B-plane

Fig 8.5 Orthogonal umbilic transition

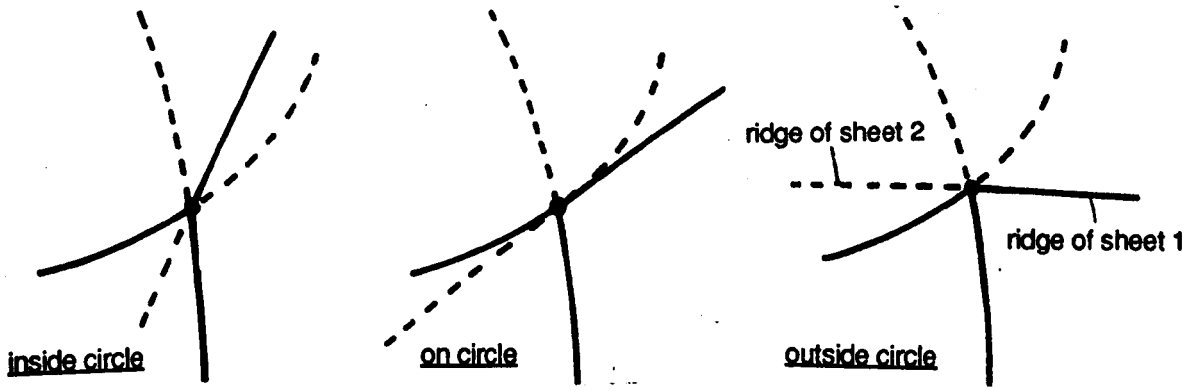


fig. 8.6) Change in pattern of half-ridges in orthogonal umbilic transition

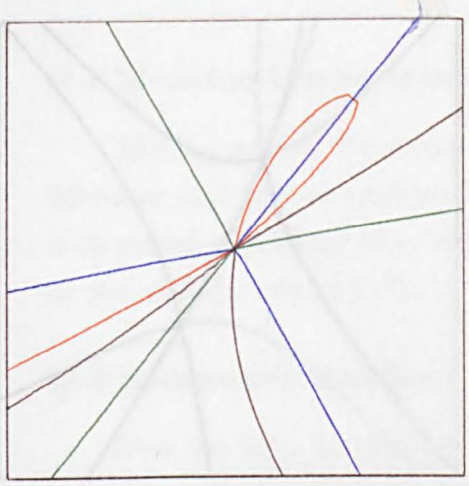
colour followed by a half ridge of the other colour. Outside the circle we have all three ridges of each colour adjacent. This latter arrangement is necessary for the parabolic umbilic transition to occur.

§8.5 Parabolic umbilic transition

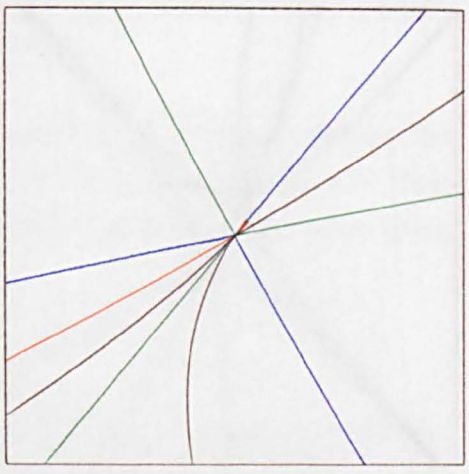
The pattern of ridges and sub-parabolic lines in the D_5 transition is shown in figure 8.7)*. Again the sub-parabolic lines are unaltered. One of the ridges undergoes the standard unfolding of a cusp. The point of self-intersection of this ridge or the cusp point lies at the umbilic. An interesting observation is that this cusp is tangent to the sub-parabolic line of the opposite sheet. Hence in both the elliptic and hyperbolic cases a ridge and a sub-parabolic line of the opposite sheet will cross.

§8.6 Birth of umbilics

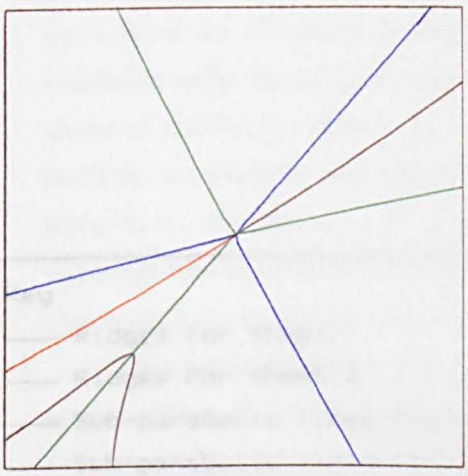
For umbilics corresponding to points on the unit circle $|\beta| = 1$ the distance squared map is not versally unfolded by the family of cubic forms. Instead it correspond to the birth of umbilics. In the generic transition two umbilics will be created out of a single point, one with a Monstar pattern of lines of curvature and the other with a Lemon pattern (fig. 8.8)*). Throughout the transition there is just one ridge which is smooth. Before the transition this is of one colour. After the transition it passes through both umbilics; between the umbilics there will be a section of ridge of the other colour. There are three sub-parabolic lines through each of these umbilics. For the star pattern there is an alternating sequence of the colours of the half sub-parabolic lines but for the Monstar the half lines of the same colour will be grouped together. During the transition the sub-parabolic lines undergo two simultaneous Morse transitions of the form $x^2 - y^2 = \alpha$ where α changes sign.



1: Elliptic Umbilic



2: Parabolic Umbilic

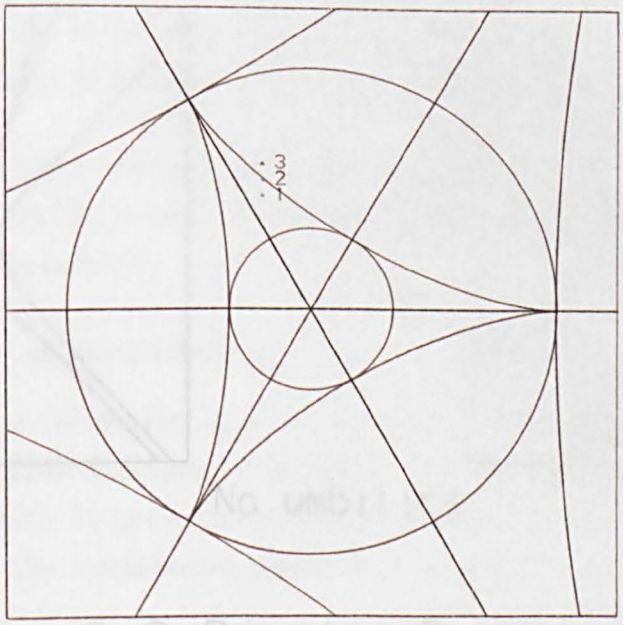


3: Hyperbolic-Star Umbilic

Key

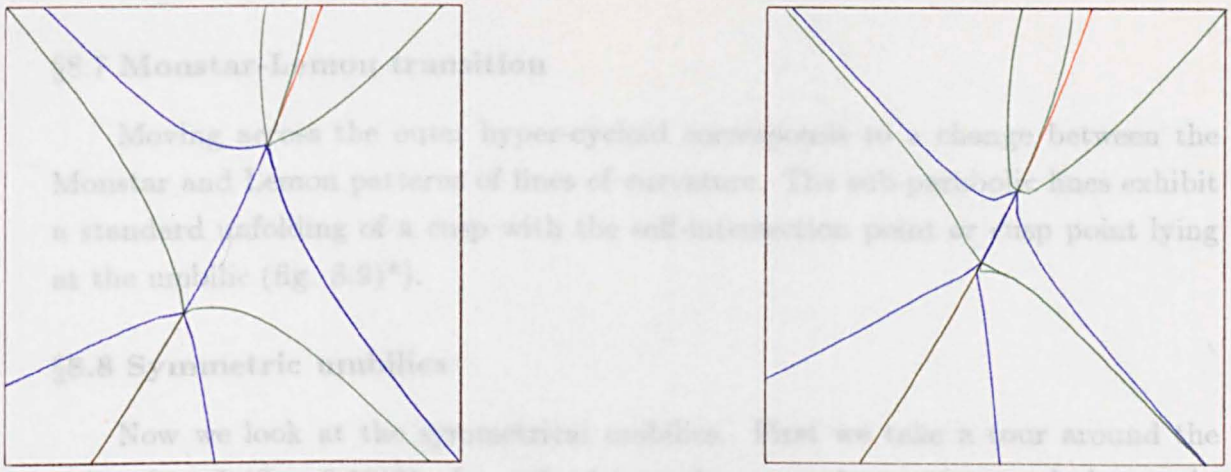
- Ridges For sheet 1
- Ridges For sheet 2
- Sub-parabolic lines For sheet 1
- Sub-parabolic lines For sheet 2

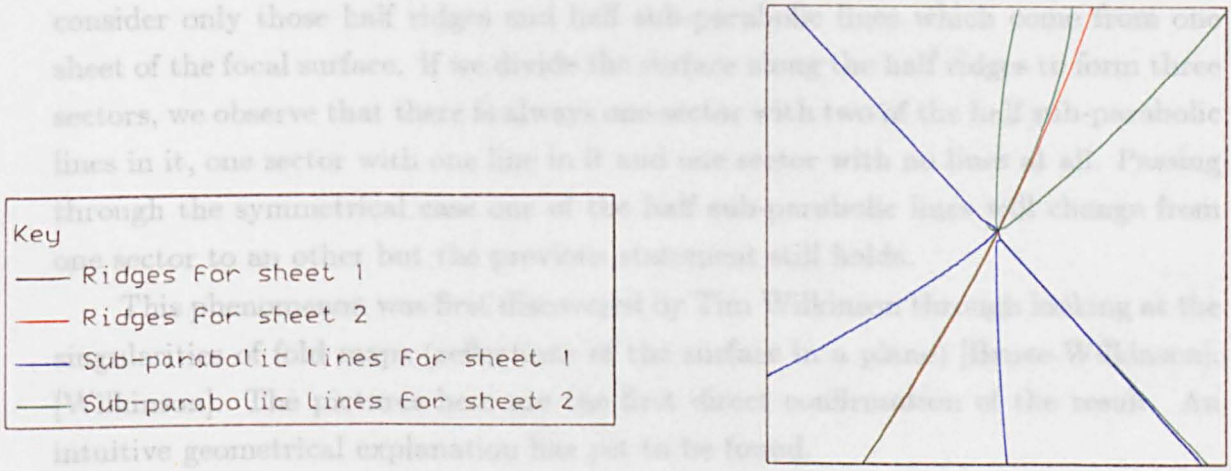
Ridges + sub-parabolic lines on surface



Position in B-plane

Fig 8.7 Parabolic umbilic transition (D5)





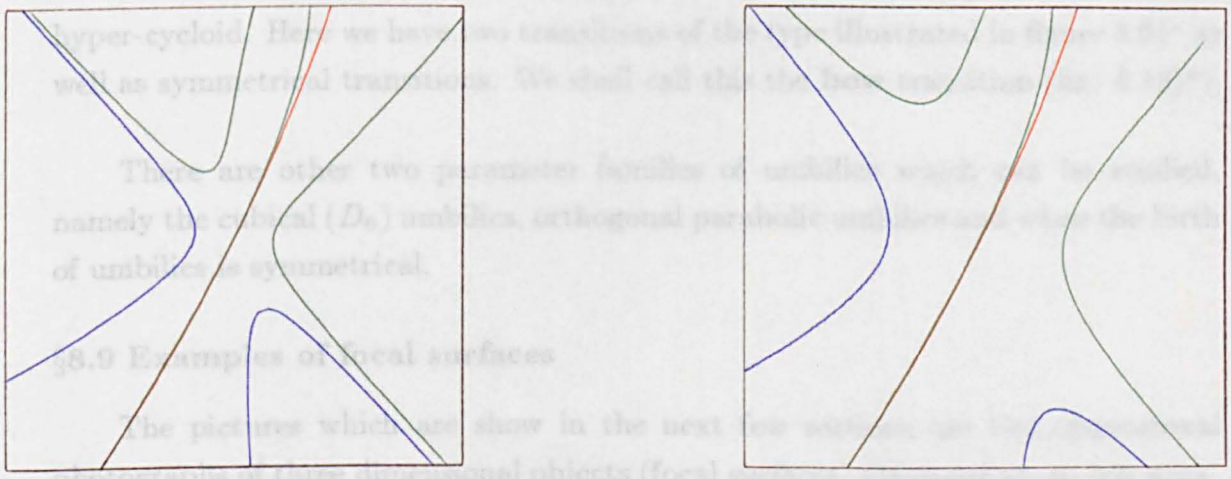


Fig 8.8 Birth of Umbilics

Ridges + sub-parabolic lines on surface

§8.7 Monstar-Lemon transition

Moving across the outer hyper-cycloid corresponds to a change between the Monstar and Lemon patterns of lines of curvature. The sub-parabolic lines exhibit a standard unfolding of a cusp with the self-intersection point or cusp point lying at the umbilic (fig. 8.9)*).

§8.8 Symmetric umbilics

Now we look at the symmetrical umbilics. First we take a tour around the point $\beta = 0$ (fig. 8.10)*). In each picture there are three ridges and three sub-parabolic lines passing through the umbilic and the half ridges or half sub-parabolic lines have an alternating sequence of colours as we move round the umbilic. Now consider only those half ridges and half sub-parabolic lines which come from one sheet of the focal surface. If we divide the surface along the half ridges to form three sectors, we observe that there is always one sector with two of the half sub-parabolic lines in it, one sector with one line in it and one sector with no lines at all. Passing through the symmetrical case one of the half sub-parabolic lines will change from one sector to an other but the previous statement still holds.

This phenomenon was first discovered by Tim Wilkinson through looking at the singularities of fold maps (reflections of the surface in a plane) [Bruce-Wilkinson], [Wilkinson]. The pictures here are the first direct confirmation of the result. An intuitive geometrical explanation has yet to be found.

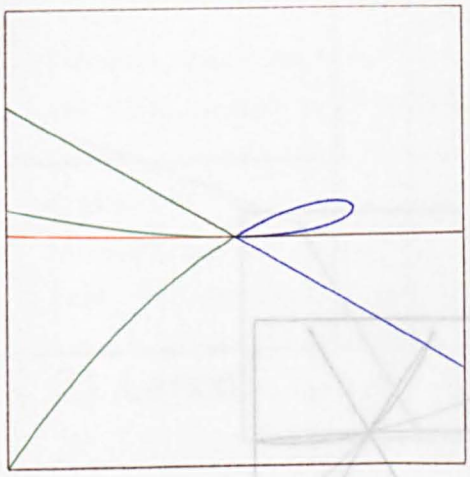
Taking a tour around the point at infinity gives figure 8.11)*. The single ridge and sub-parabolic line become tangent each time we have a symmetrical umbilic.

Another interesting transition happens if we study the cusp point on the outer hyper-cycloid. Here we have two transitions of the type illustrated in figure 8.9)* as well as symmetrical transitions. We shall call this the **bow** transition (fig. 8.12)*).

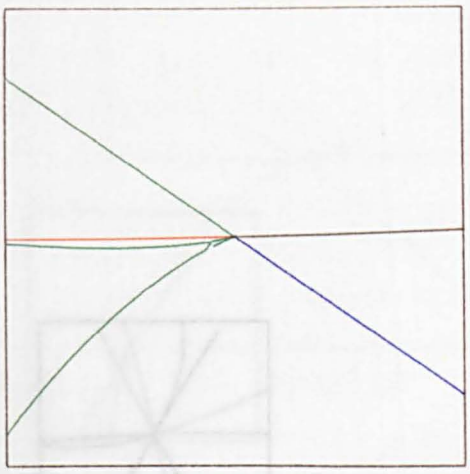
There are other two parameter families of umbilics which can be studied, namely the cubical (D_6) umbilics, orthogonal parabolic umbilics and when the birth of umbilics is symmetrical.

§8.9 Examples of focal surfaces

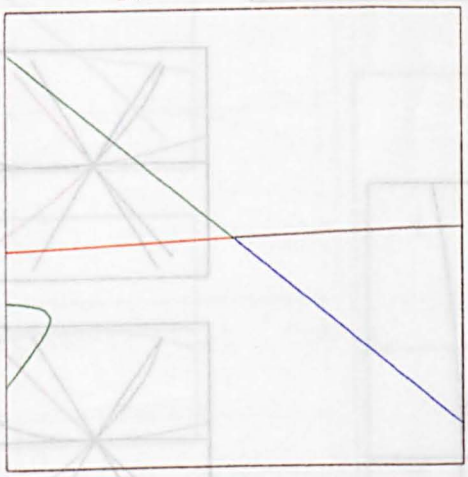
The pictures which are show in the next few sections are two dimensional photographs of three dimensional objects (focal surfaces) displayed on an Iris workstation (see §9.8.2). To get a three dimensional feal for these object it is preferable to see them on the workstation where it is possible to rotate them in real time.



1: Hypebolic - Monstar



2: Moment of Transition

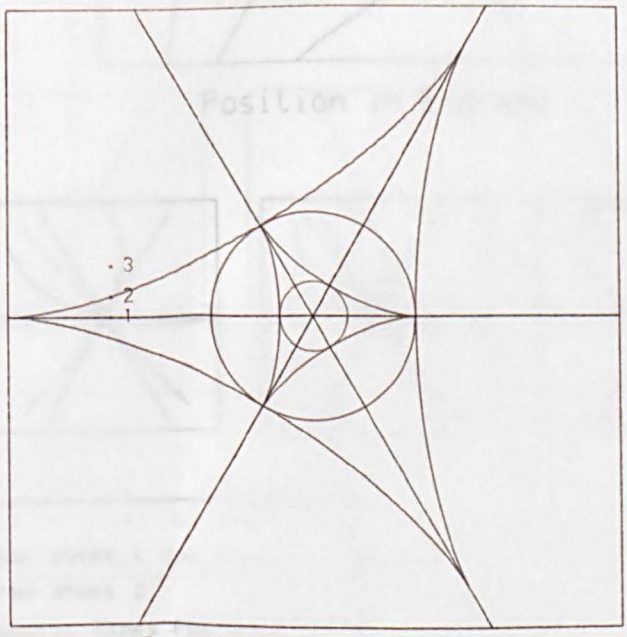


3: Lemon

Key

- Ridges For sheet 1
- Ridges For sheet 2
- Sub-parabolic lines For sheet 1
- Sub-parabolic lines For sheet 2

Ridges + sub-parabolic lines on surface



Position in B-plane

Fig 8.9 Monstar - Lemon transition

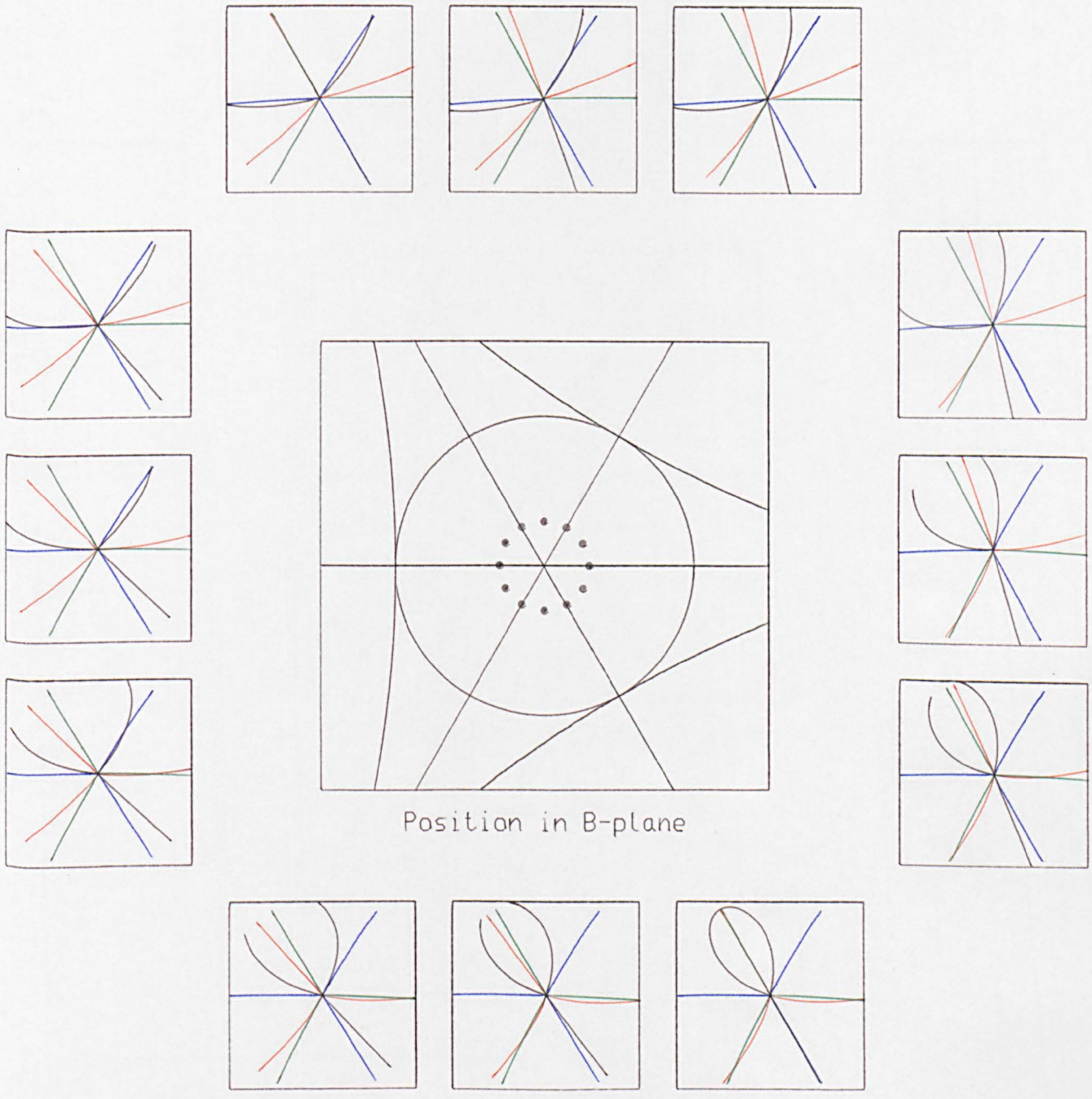


Fig 8.10 An orbit of the point 0,0

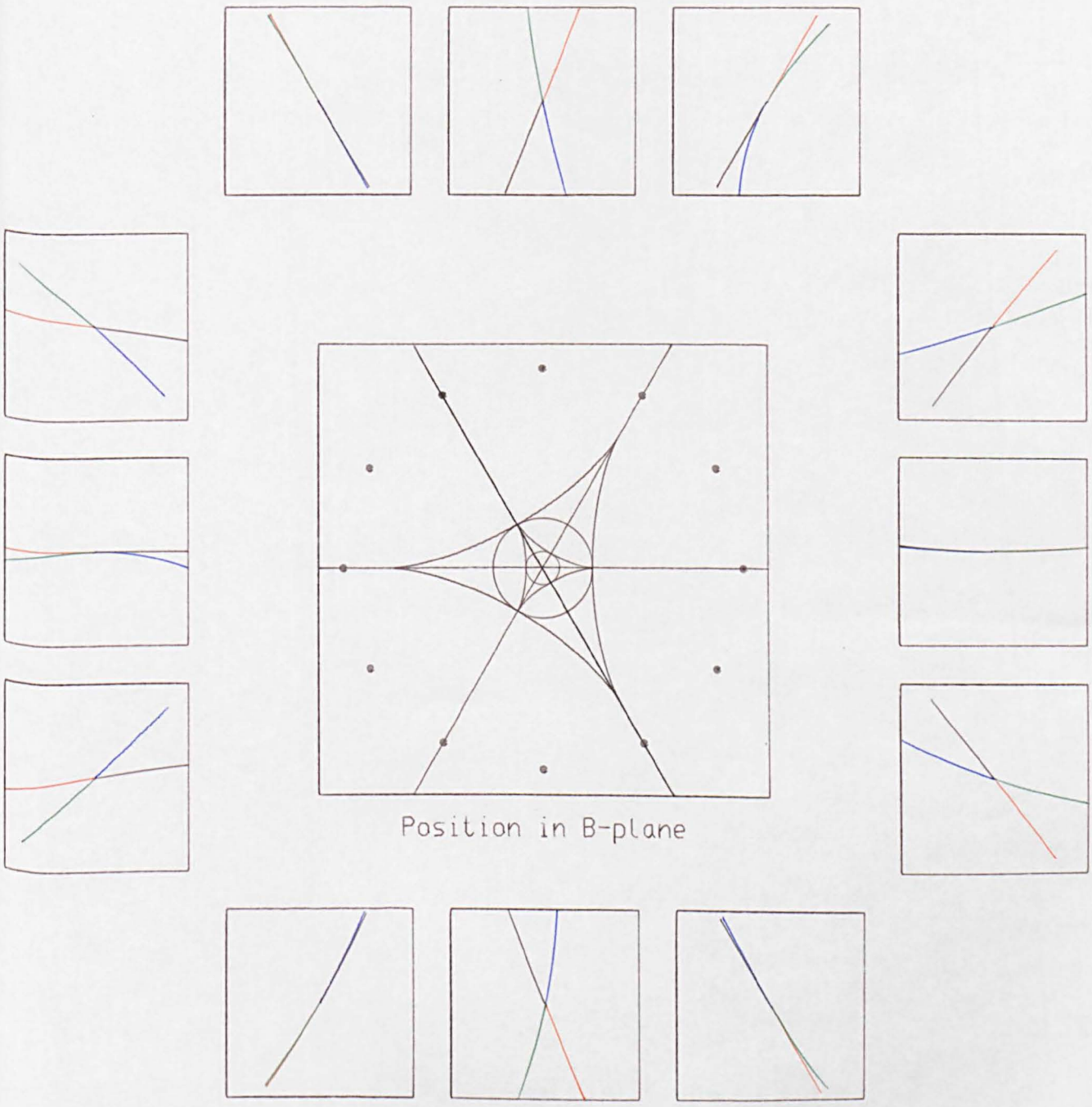
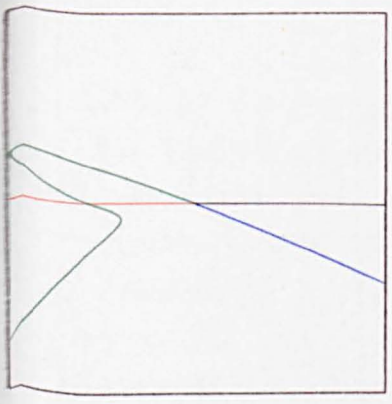
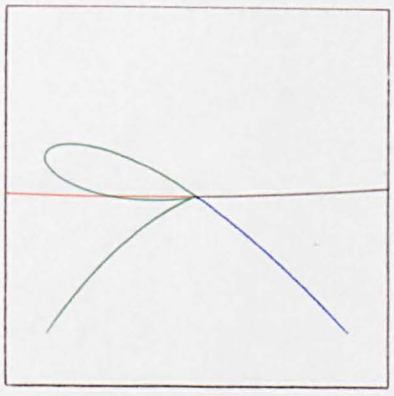


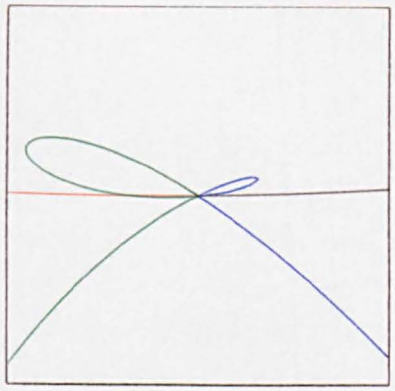
Fig 8.11 An orbit of the point at infinity



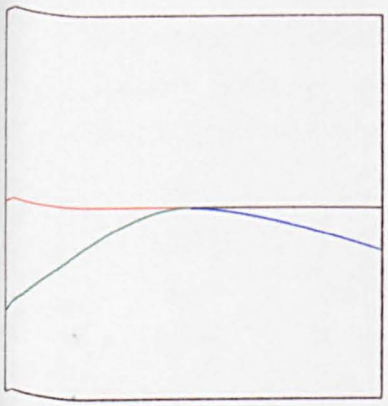
7: Lemon



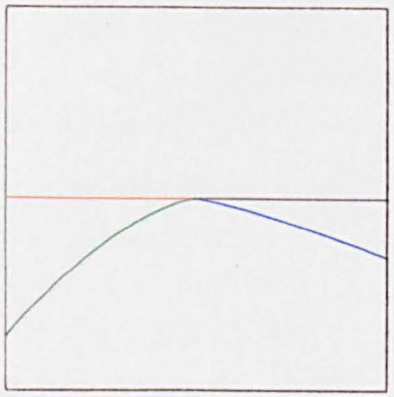
8: On outer tri-cusp



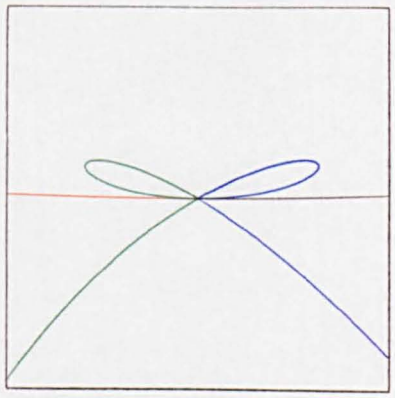
9: Monstar



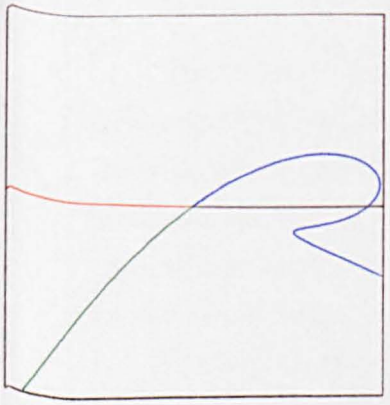
: Symmetrical Lemon



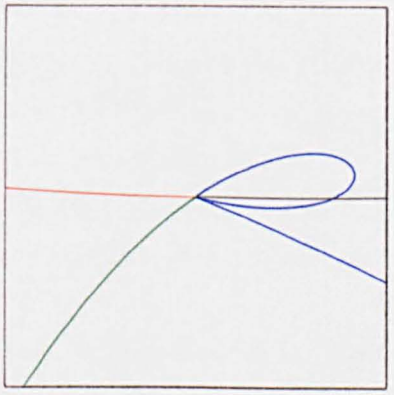
1: On cusp



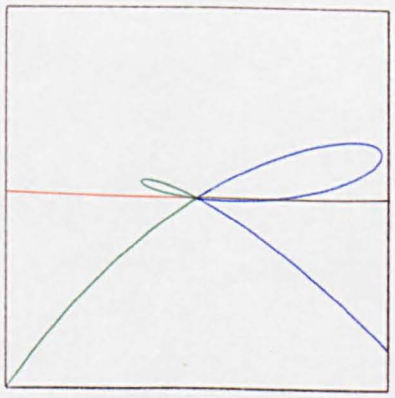
2: Symmetrical Monstar



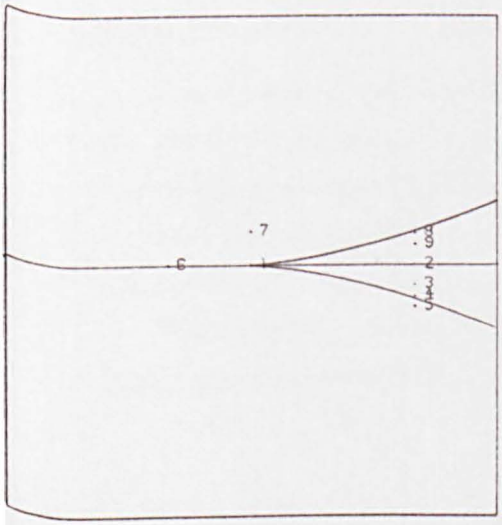
5: Lemon



4: On outer tri-cusp



3: Monstar



Key

Ridges For sheet 1

Ridges For sheet 2

Sub-parabolic lines For sheet 1

Sub-parabolic lines For sheet 2

Fig 8.12 The Bow transition

Because the workstation only arrived after the end of my Ph.D. grant there has not been enough time to fully developed these pictures and several problems still remain. One problem is that near the ribs the surfaces as drawn are only approximations to their actual shapes. In particular the surfaces as draw do not come up to touch the rib lines (fig. 8.**a)*). They should actually form a cuspidal edges here. The reasons for this error is discussed in §9.12. If two adjacent surfaces are close together in space then the graphics software can not tell which is the nearest this may lead to mis-colouring as can be seen in fig. 8.**c)*.

The colours have been chosen to give a general indication of the Gaussian curvature and should not be taken as precise values. Furthermore they do not vary linearly with the value of the curvature. The method of colouring is discussed in §9.12. The colours are coded as follows:

Orange - large positive Gaussian curvature,
Green - positive Gaussian curvature,
Red - negative Gaussian curvature,
Blue - large negative Gaussian curvature.

The ribs and parabolic lines are also drawn these are coloured as follows:

White - ribs,
Yellow - parabolic lines.

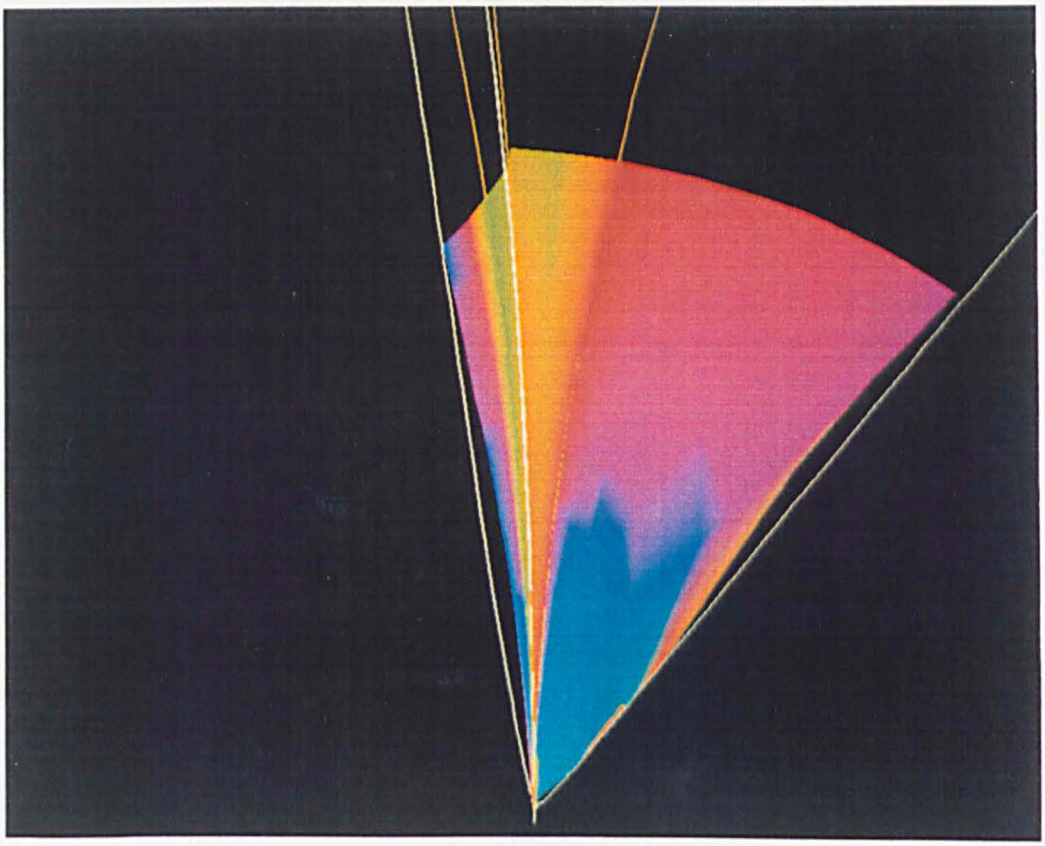
The colours have only been defined at a finite number of data points and between these points the colours are chosen by interpolation. This has lead to the layering effect seen in figure 8.**. Across the ribs or parabolic lines there should be a sudden change in colour (from positive to negative curvature). Unfortunately the data points have not been chosen to lie on these curves so the colours tend to spill across these lines.

Work is now in progress to improve these figures. The reader is referred to sections 9.8–9.12 where the production of the picture is discussed in detail.

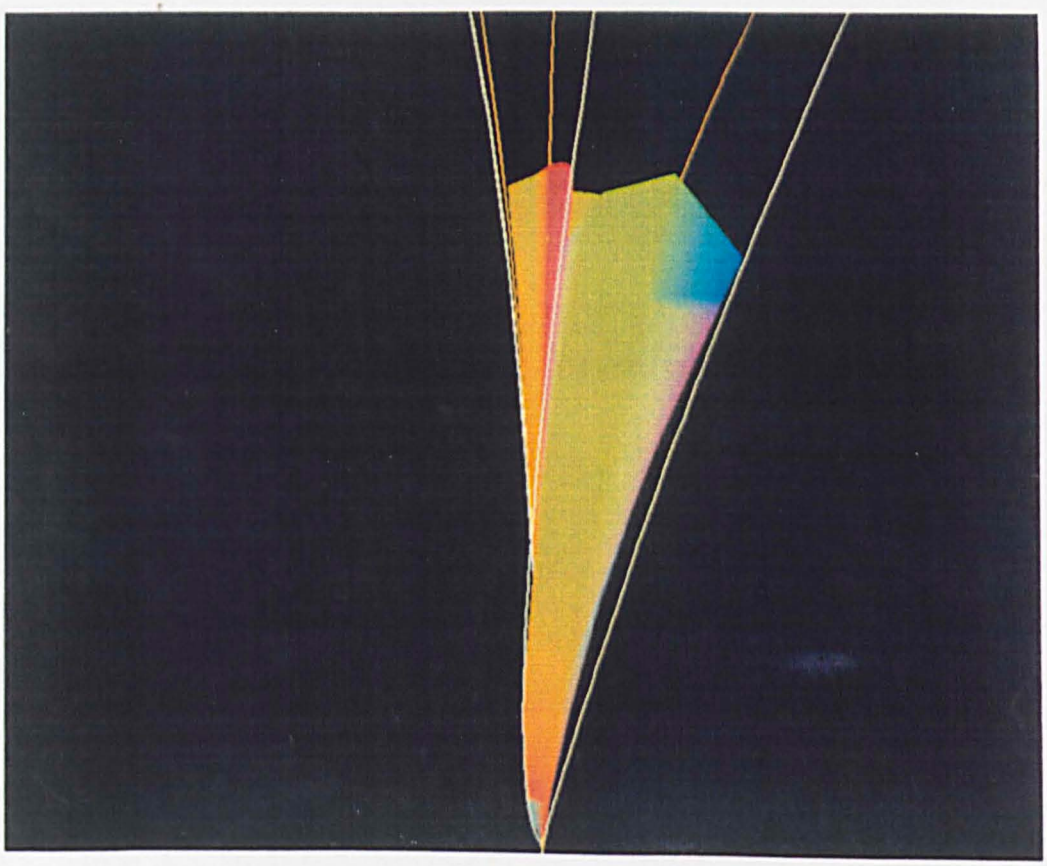
§8.10 Parabolic lines on an elliptic umbilic

The standard pictures (fig. 6.8) of elliptic and hyperbolic umbilics do not show any parabolic lines and only recently was their presence known [Bruce-Wilkinson]. One of the main motivations behind this project was to deduce how parabolic lines occur on a focal surface. In particular it was not known how the regions of positive and negative Gaussian curvature were arranged.

We will just look at one sheet of the focal surface. In the standard model of an elliptic umbilic the whole surface has positive Gaussian curvature and each cuspidal

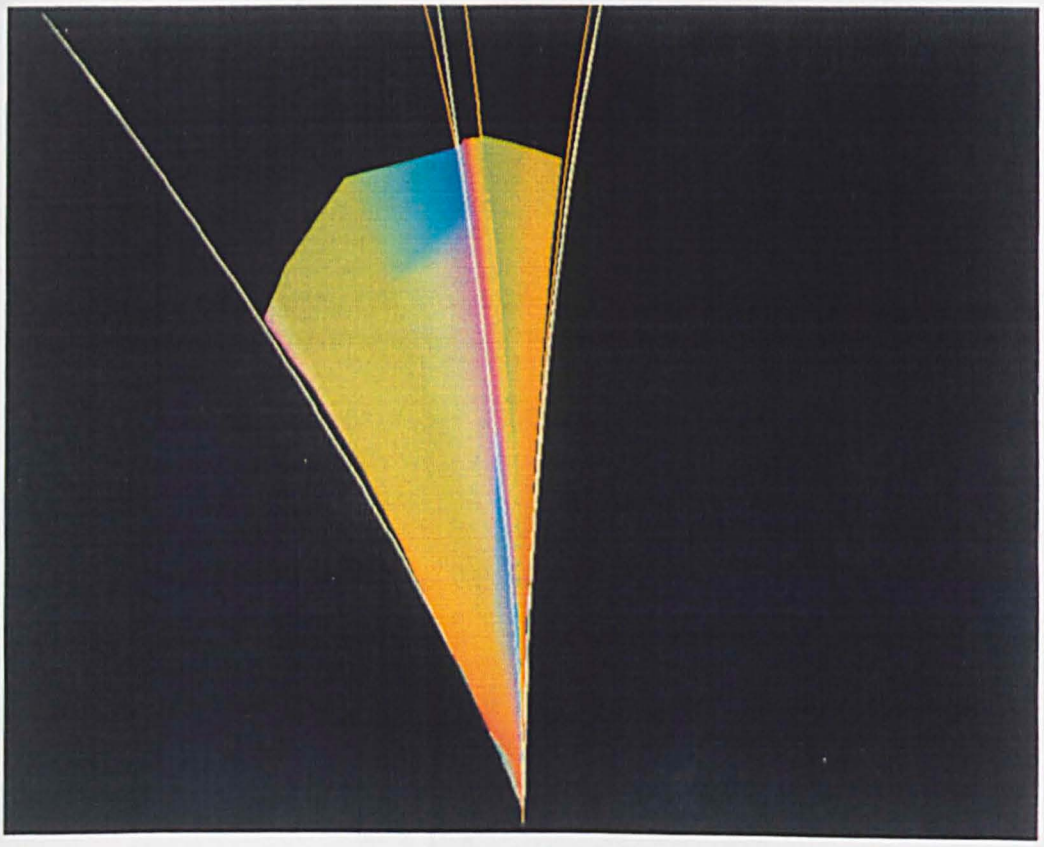


a)

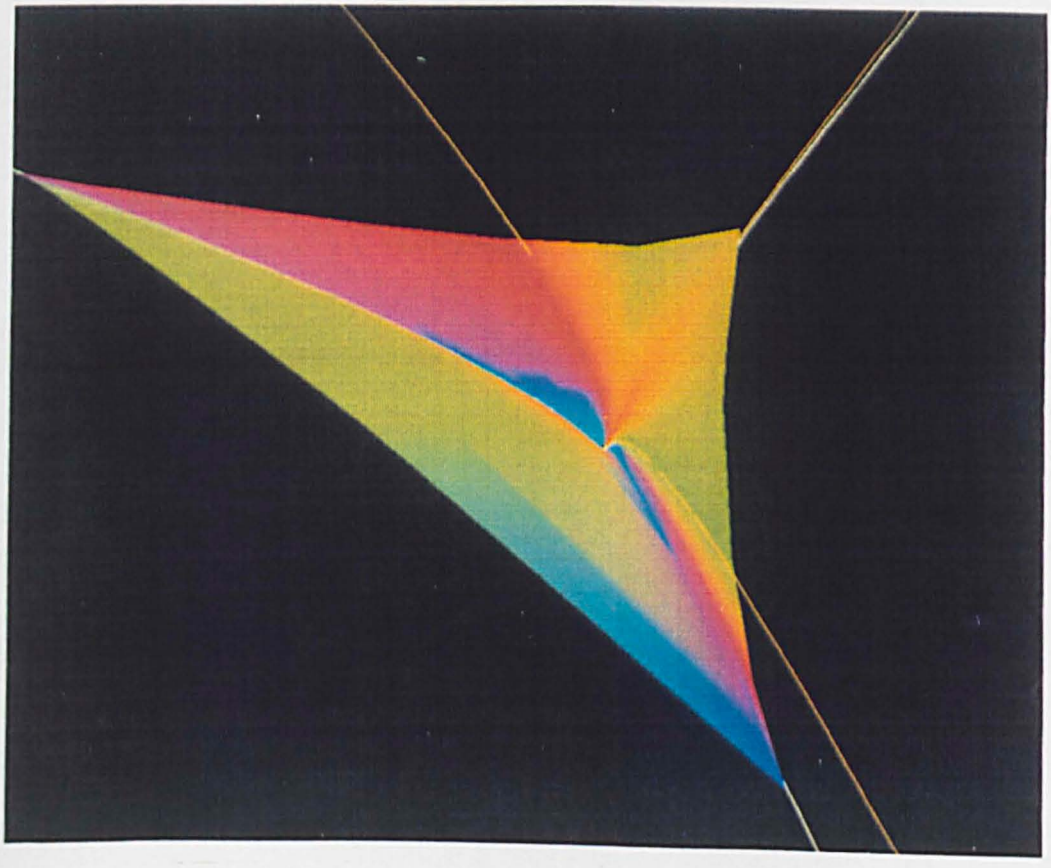


b)

fig 8.14) One sheet of an elliptic umbilic focal surface.



c)



d) View from above.

fig 8.14) cont. One sheet of an elliptic umbilic focal surface.

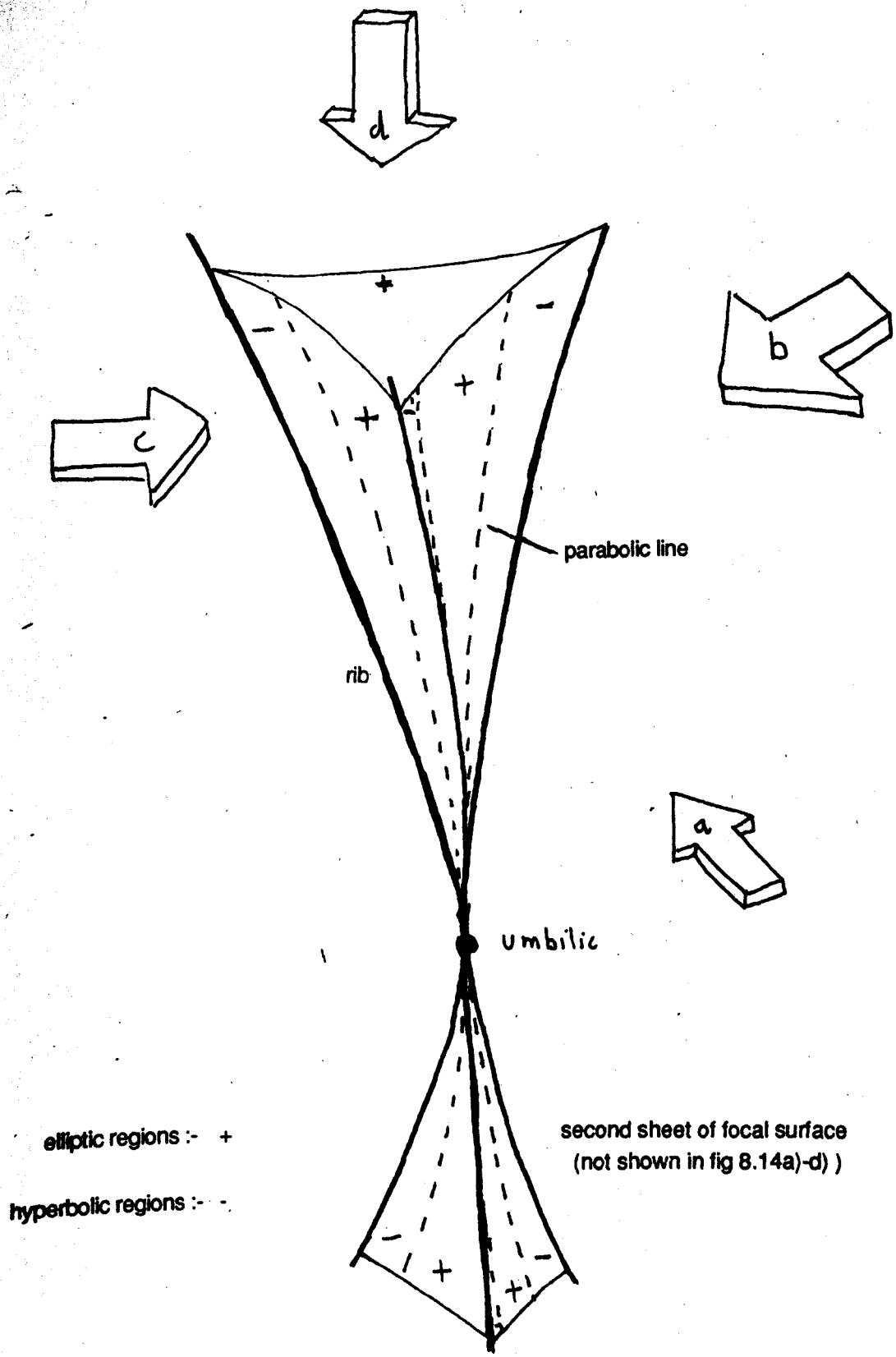


fig 8.14e) Sketch of the elliptic umbilic shown in fig 8.14a)-d) indicating the directions of the different views

edge lies in one vertical plane. If we allow one of the edges to bend so that it leaves this plane (the quantity κ of §7.4 becomes non-zero), then the curvature on opposite sides of the cuspidal edge will have opposite signs thus creating a parabolic line and a region of negative Gaussian curvature (fig. 8.13). Generically we would expect all

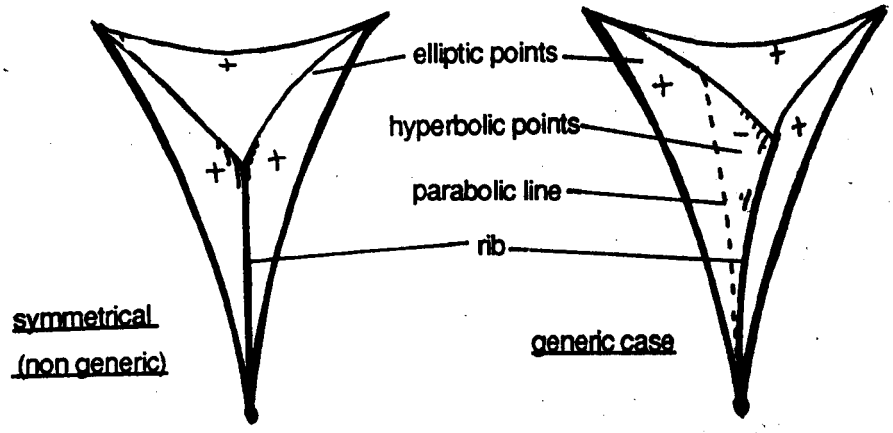


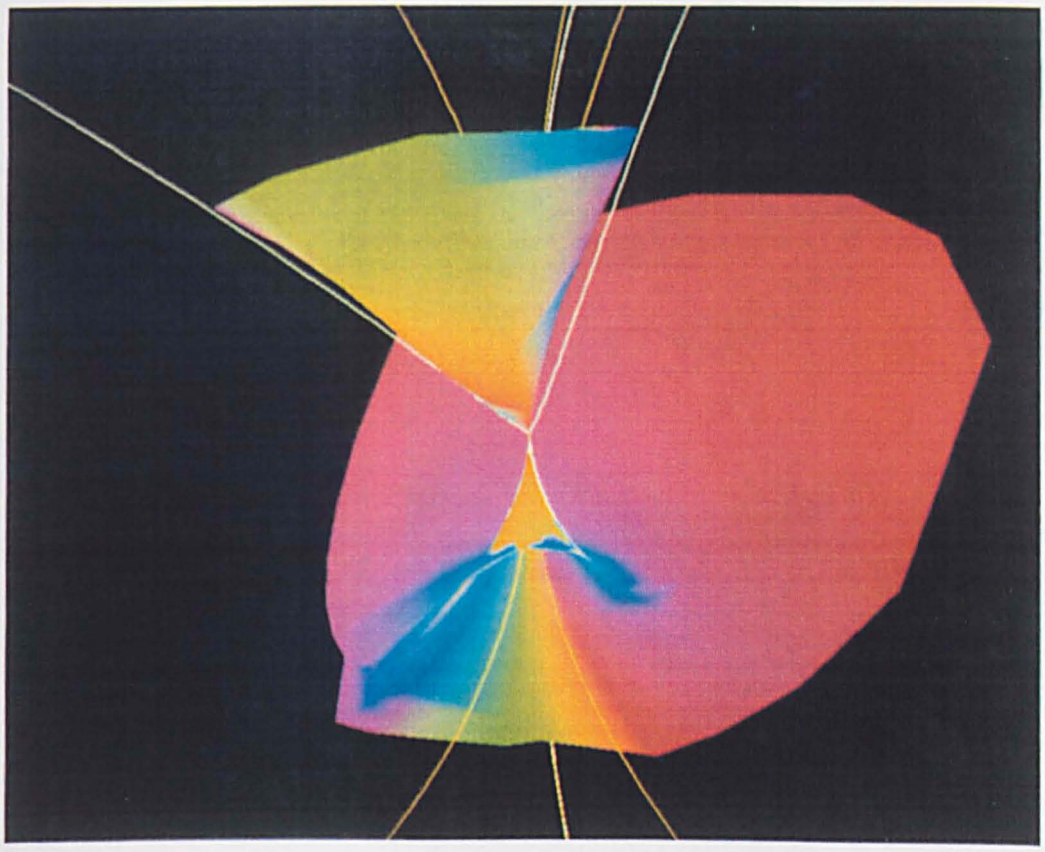
fig. 8.13) Creation of a region of hyperbolic points on an elliptic umbilic focal surface

three edges to be curved in this way and hence there will be three parabolic lines and three hyperbolic regions. However if we cut the focal sheet up into three pieces along the ribs (as in §8.8) we see that there is one piece with no parabolic lines and consisting entirely of elliptical points. Another piece will have one parabolic line and the last piece has two parabolic lines with an elliptic region between two hyperbolic regions. In the example shown (fig. 8.14)*) we are near the symmetrical case where one rib and one parabolic line would be identical but we can still see the way in which the other two ribs curve. Figure 8.14e) is a sketch of the umbilic which shows the elliptic and hyperbolic regions.

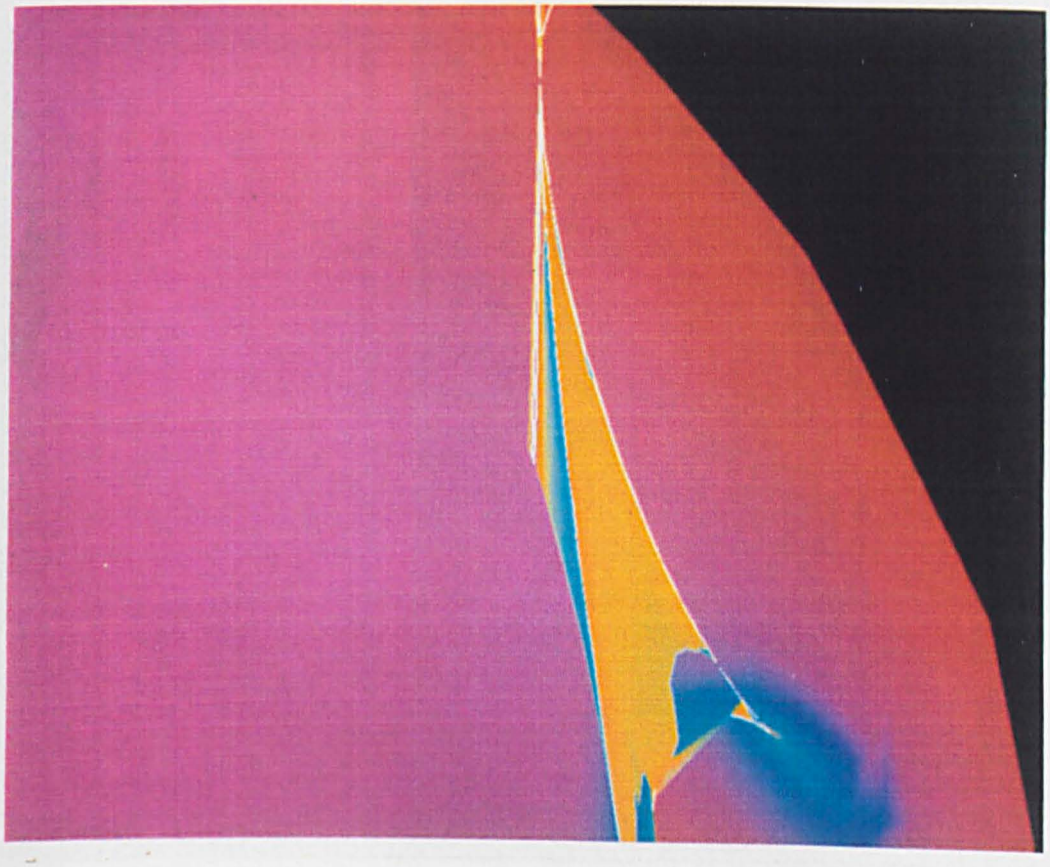
§8.11 Parabolic Umbilic transition

One of the most interesting transitions on a focal surface is the parabolic umbilic transition, where the type of umbilic changes from being elliptic to being an hyperbolic star. In this transition the number of cuspidal edges on the focal surface changes from three to one but the number of parabolic lines remains unchanged. We can think of this transition as having three phases: elliptic, hyperbolic and parabolic corresponding to the type of umbilic. We also number the sheets 1 and 2 for ease of reference.

As we have seen in §8.5 in the elliptic phase one of the ridges, of sheet 1 say, forms a small loop from the umbilic back to itself (fig. 8.15). There are two A_4 points on this loop. On sheet 1 of the focal surface one of the ribs will form a triangle with one point at the umbilic and two swallowtail points (fig. 8.17f)*). There will

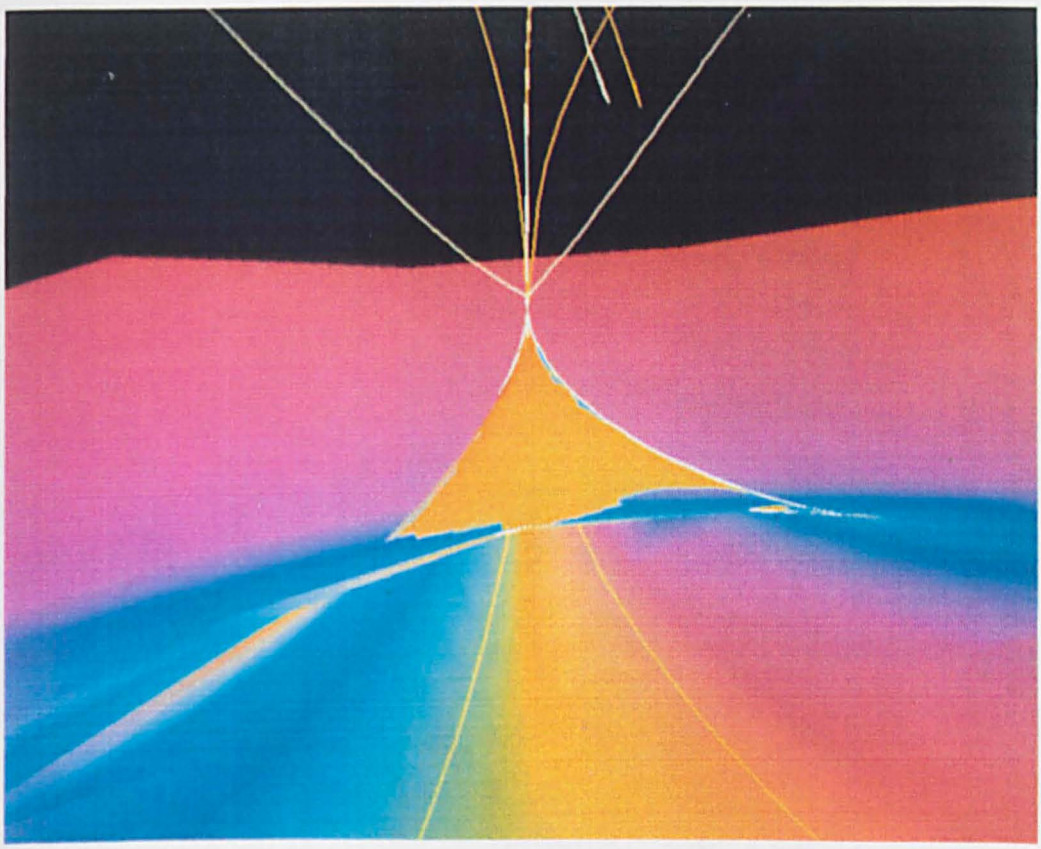


a) both sheets

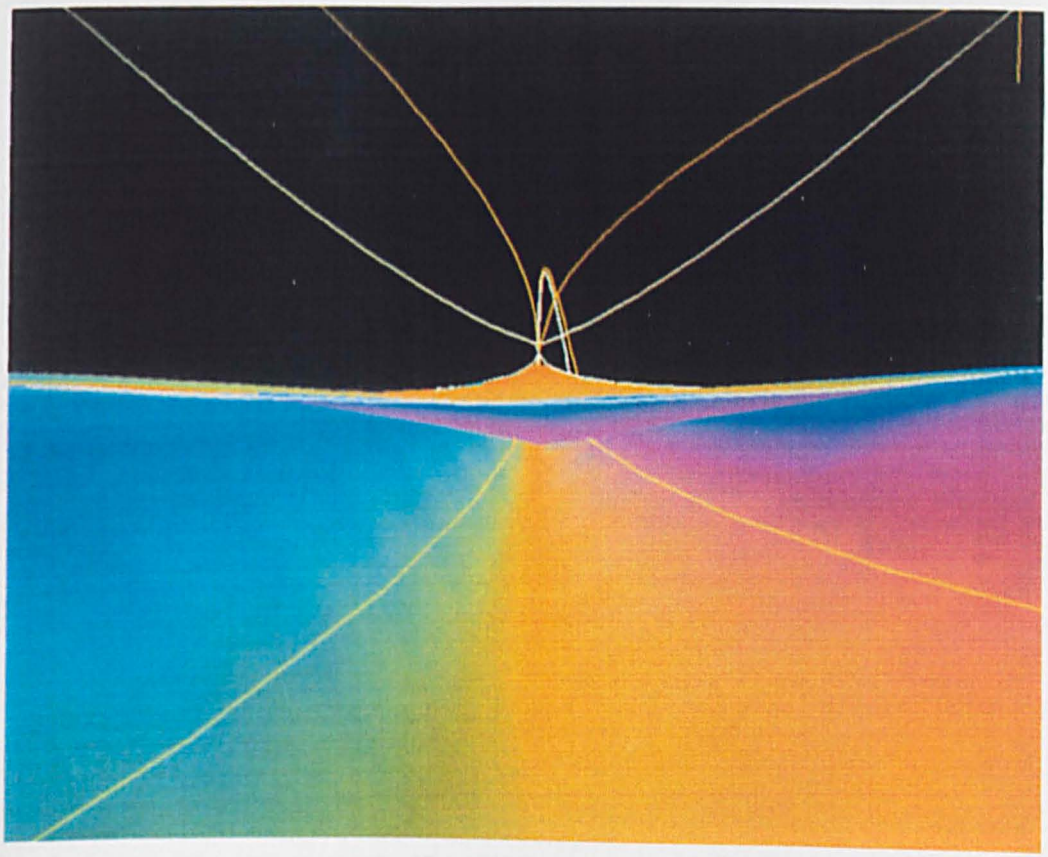


b) close up of umbilic centre

fig 8.17) Elliptic phase of parabolic umbilic transition.

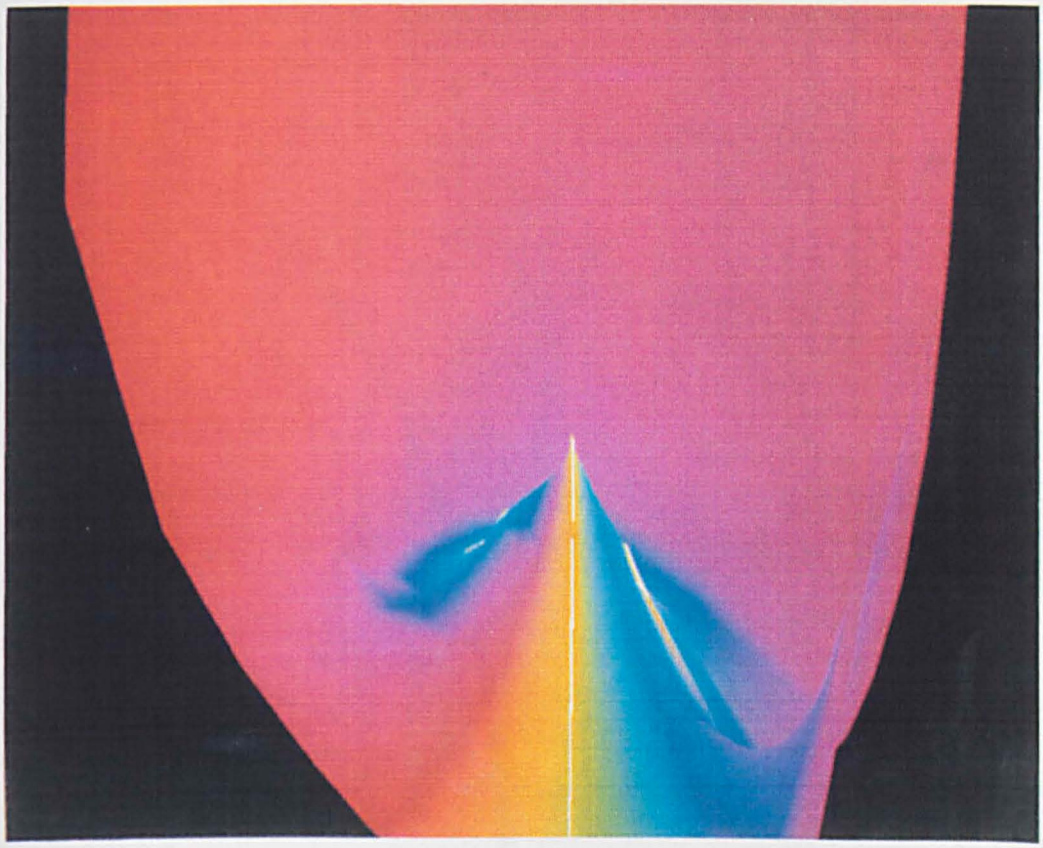


c) view of focal sheet 1 from below

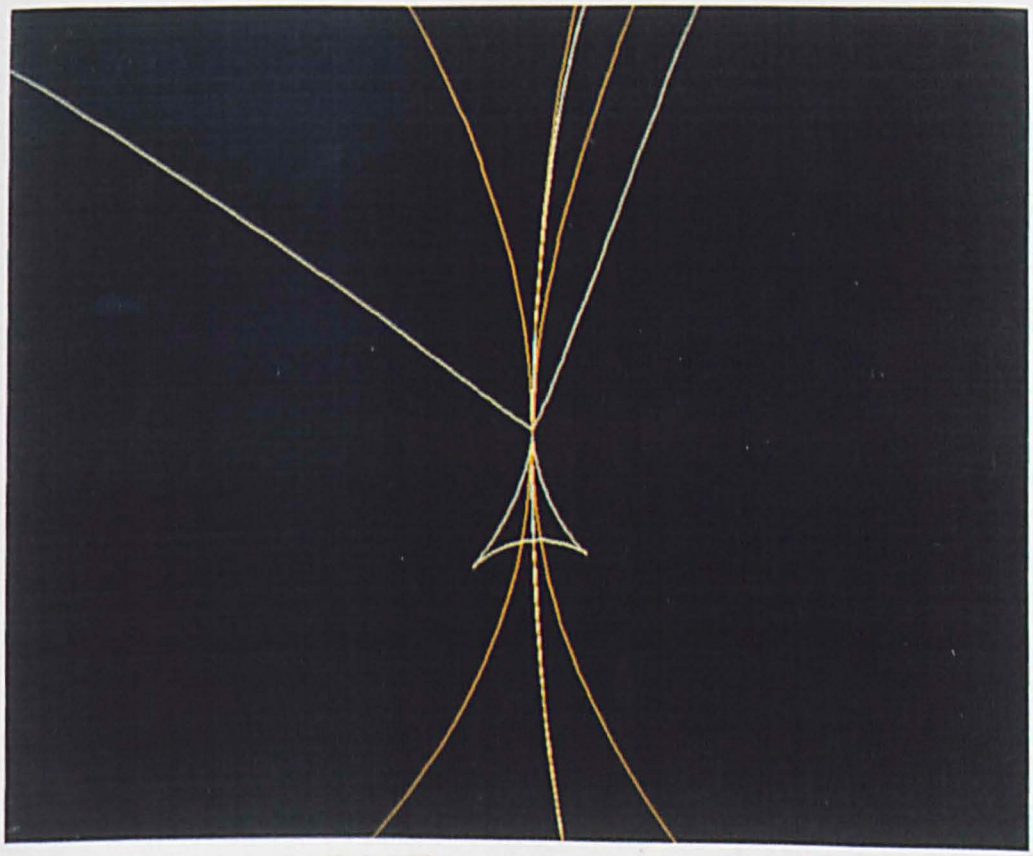


d) view from below showing self intersection

fig 8.17) cont. Elliptic phase of parabolic umbilic transition.



e) view of focal sheet 1 from behind



f) the ribs and parabolic lines of both sheets

fig 8.17) cont. Elliptic phase of parabolic umbilic transition.

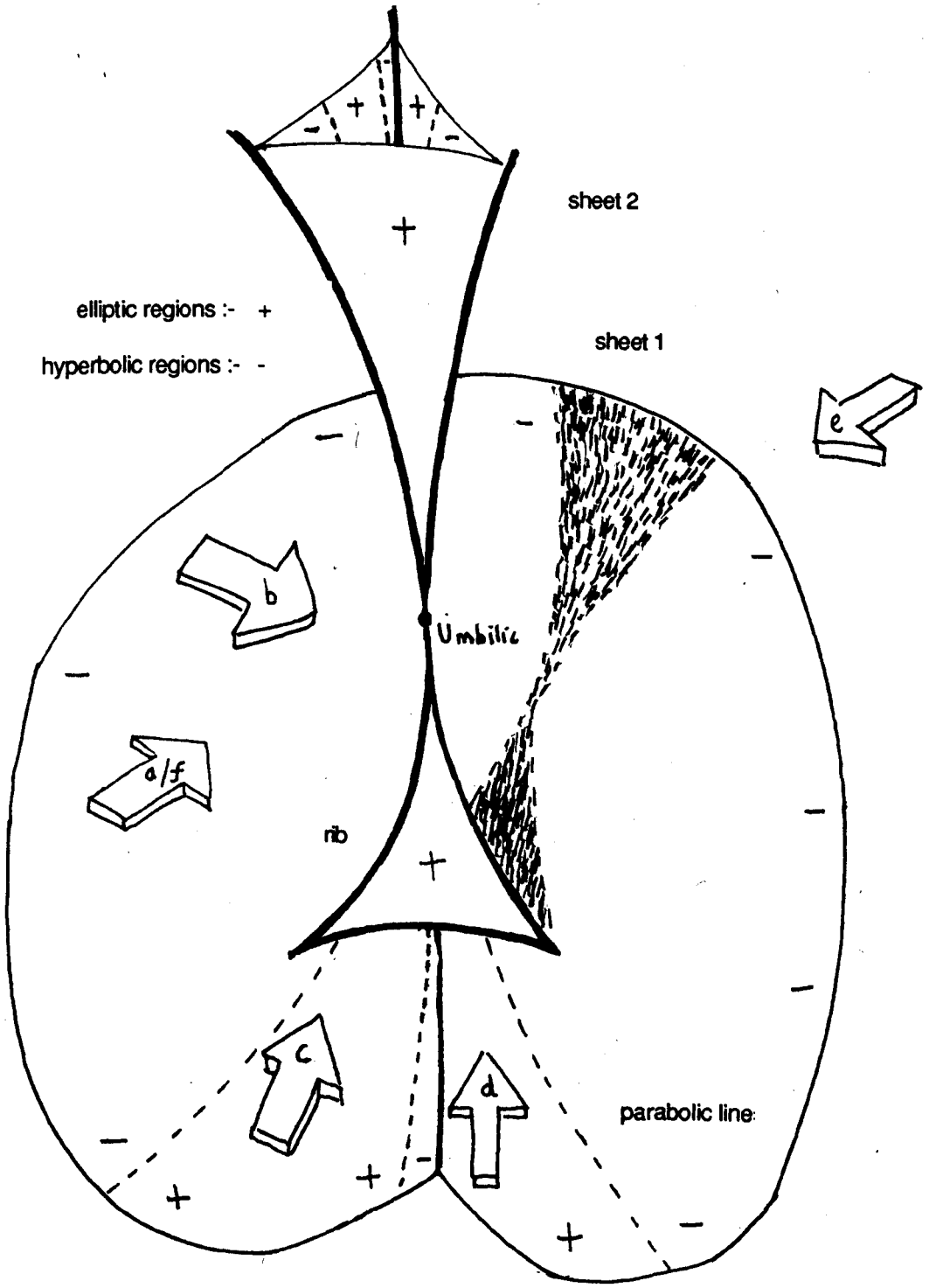


fig 8.17g) Sketch of the umbilic shown in fig 8.17a)-f) indicating the directions of the different views.

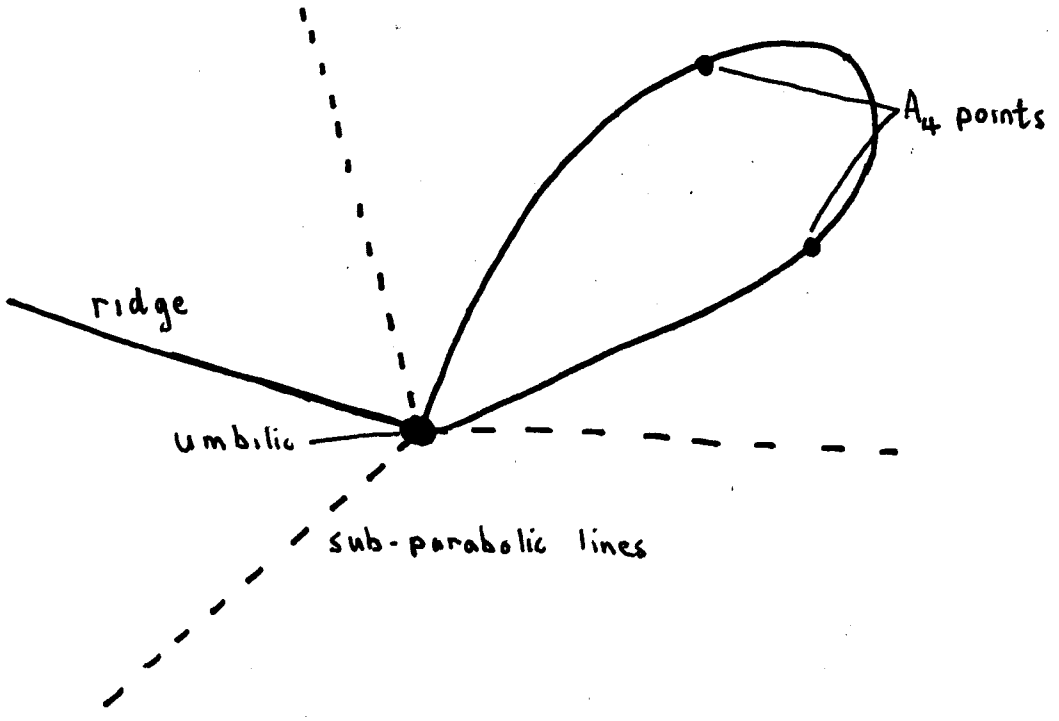


fig 8.15) The ridges and sub-parabolic lines of sheet 1 in elliptic phase of D_5 transition.

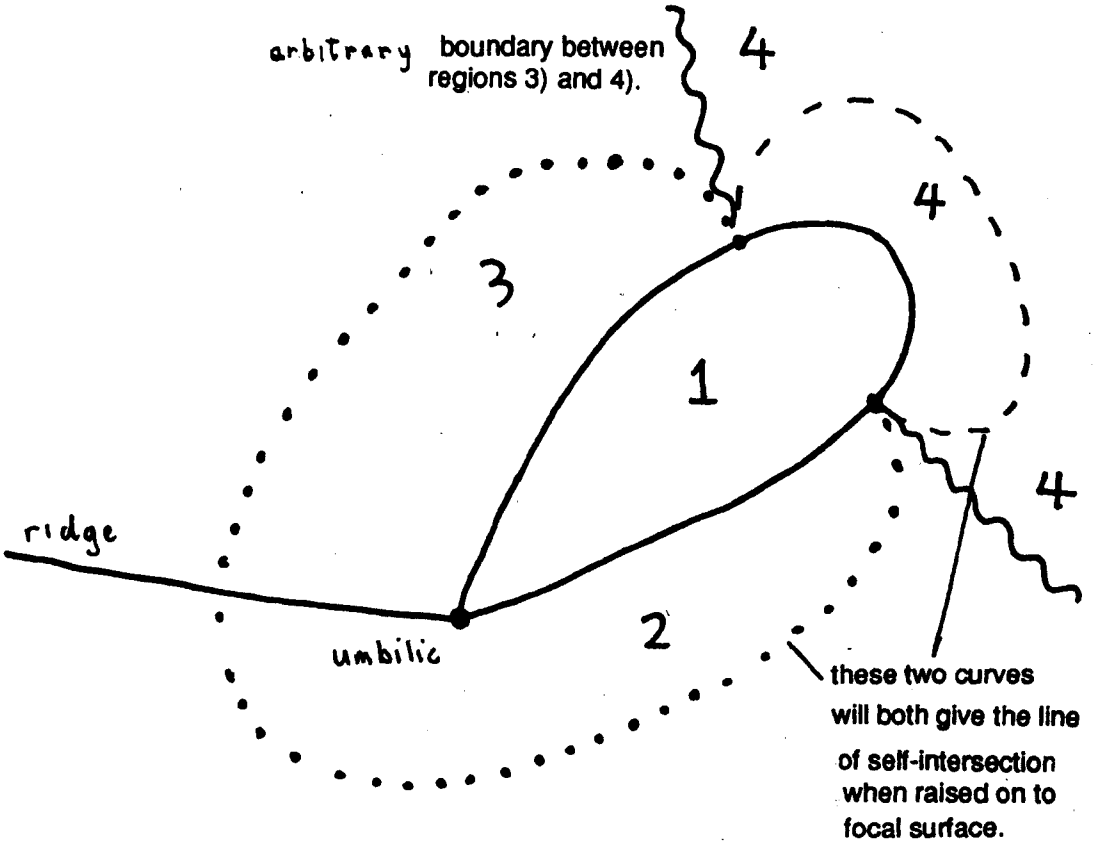


fig 8.16a) Regions 1) - 4) on original surface, in elliptic phase of D_5 transition.

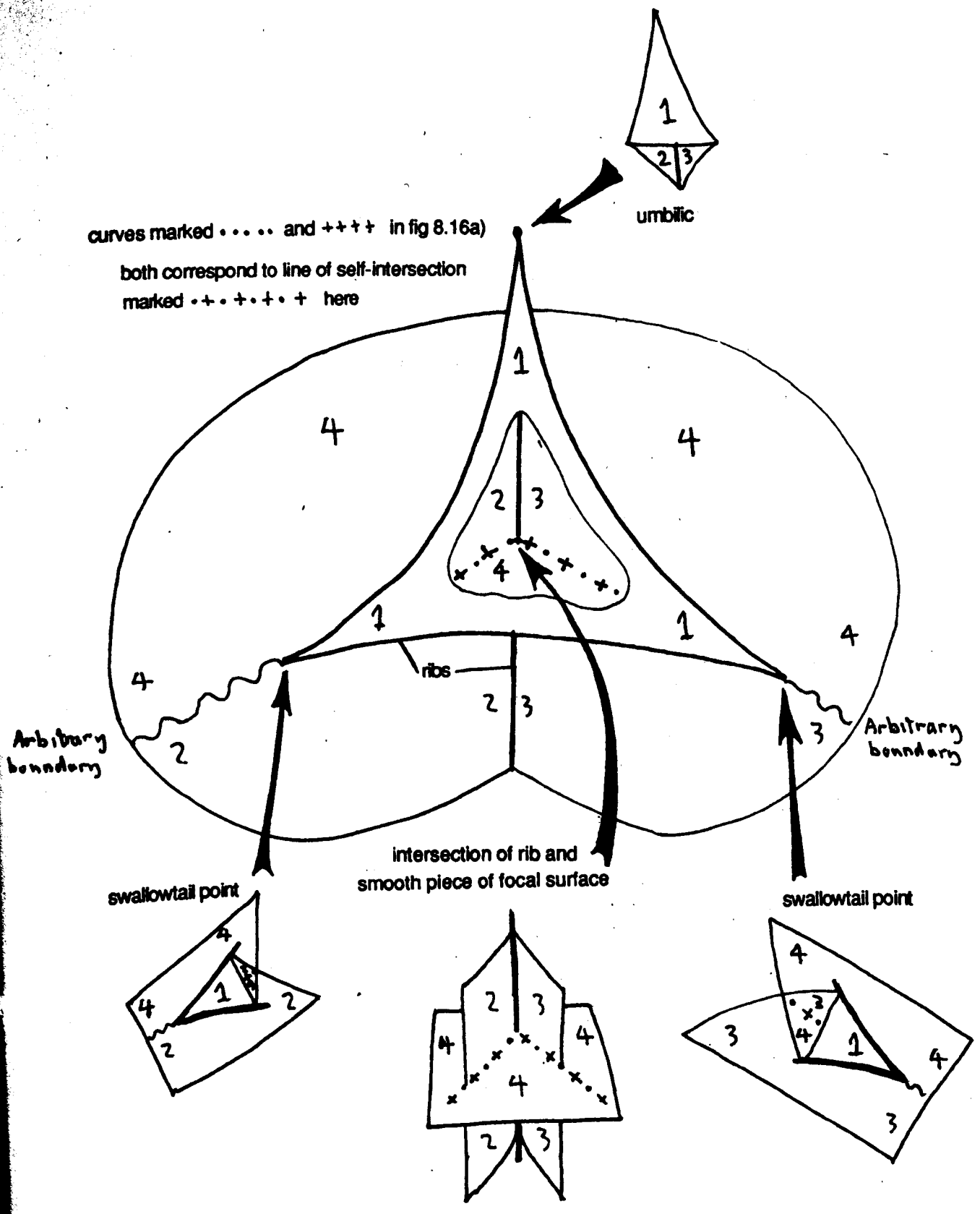


fig 8.16b) Cut away view of sheet 1 of a focal surface in elliptic phase of D_5 transition, showing regions 1)-4) corresponding to those in fig8.16a).

also be one non-singular rib going straight through the umbilic. On sheet 2 we will have the normal arrangement of ribs and parabolic lines which we would expect near an elliptical umbilic. The correspondence between the regions of the surface and the regions of one of the focal surfaces is shown in figures 8.16a) and b). Much of the surface in figure 8.16 has been removed to help illuminate the geometry.

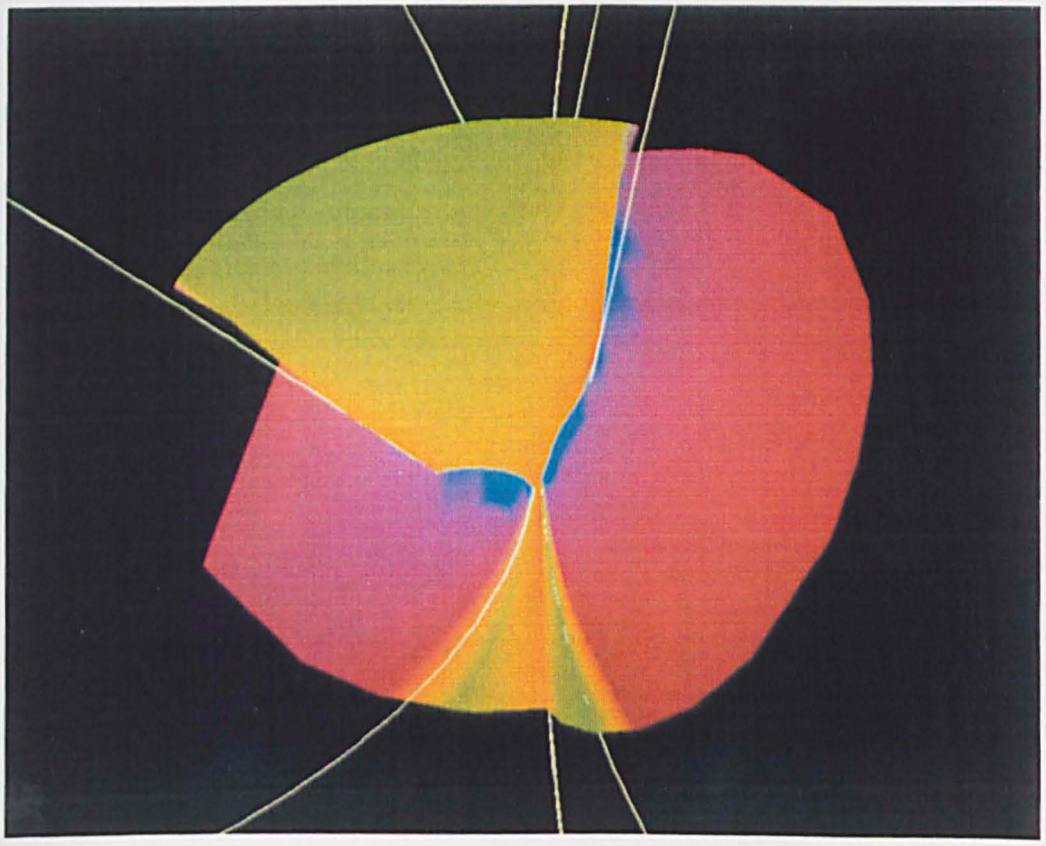
In figure 8.16a), b) we have devided the surface and sheet 1 of the focal surface into two four regions where points in region 1 in of the surface correspond to points in region 1 of the focal surface. These regions have boundarys along the ribs/rigdgs. We have also introduced two artificial boundarys, starting at A_4 points and going off to infinity, to help distinguish between regions 2 and 4 and regions 3 and 4.

From a great distance sheet 1 resembles a sheet of an hyperbolic umbilic with just one rib line. Closer up this sheet looks like an elliptic umbilic with the familiar triangular pyramid (fig. 8.17b)*). Figure 8.17)* show various views of such a surface. In fig 8.17c–d)* a rib and a parabolic line should run down the middle of the surface. These curves are hidden from view in these pictures. However the edge can be seen from behind the surface in figure 8.17e). In figures 8.17 b), d), e) we can see the line of self-intersection which runs between the two A_4 points and crosses one of the ribs. Figure 8.17g) is a sketch of the umbilic which shows the elliptic and hyperbolic regions.

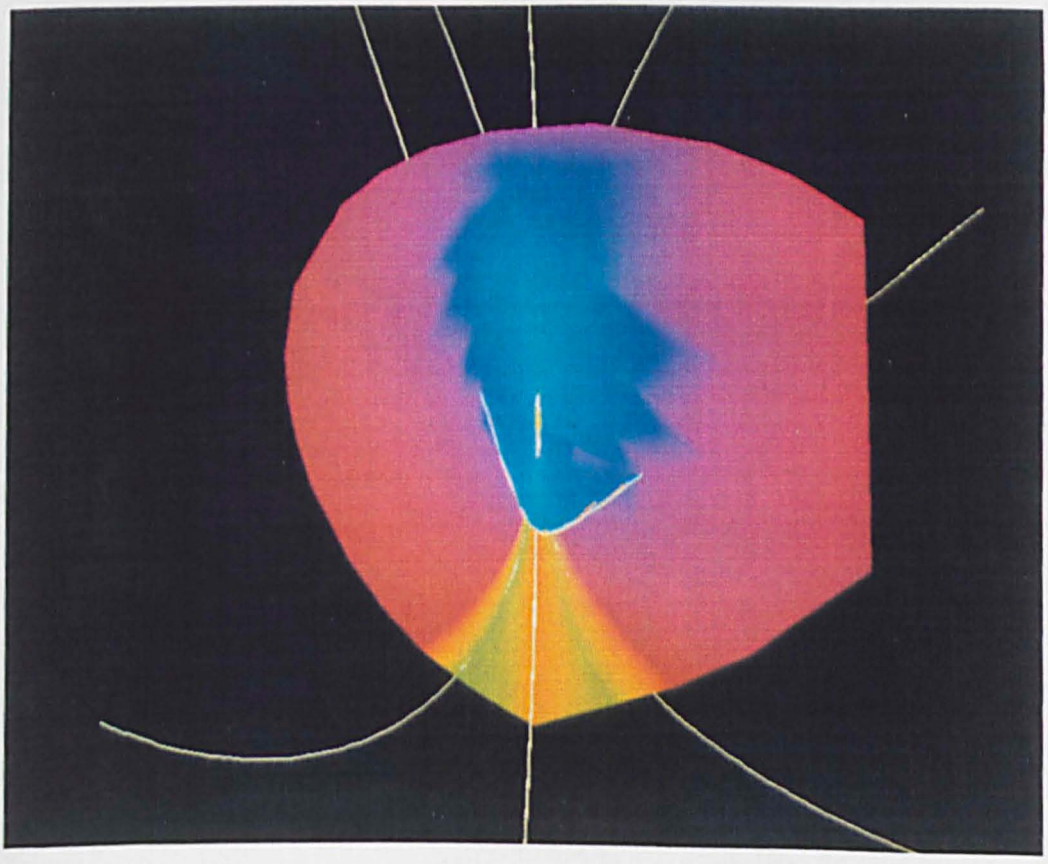
As the surface changes towards the parabolic phase the triangle shrinks down to a point. At the moment of transition sheet 1 resembles a hyperbolic umbilic and sheet 2 resembles an elliptical one. However this is only approximately true as the ribs do not pass smoothly through the umbilic.

In the hyperbolic phase two of the ribs of sheet 2 join together to form a smooth cuspidal edge which moves away from the umbilic (fig. 8.18f)*). In this example both ribs and parabolic lines have been drawn in white. This forms a pocket with three cuspidal edges at the top, one of which runs down to the umbilic (fig. 8.18d)*). The two sheets intersect with the line of intersection running round the pocket, starting and ending at the umbilic (fig. 8.18 a), b)). The umbilic is obscured in fig. 8.18a) but can be seen in the middle of 8.18b). From a distance sheet 2 resembles an elliptic umbilic but close up we have a standard hyperbolic umbilic. Throughout this phase of the transition sheet 1 looks like a hyperbolic umbilic at all scales (fig. 8.18c)). Figure 8.18g) is a sketch of the umbilic which shows the elliptic and hyperbolic regions. The indivdual sheets of the umbilic are shown in figure 8.19).

The presence of parabolic lines do not complicate the transition as they do not lie in the areas which change most.

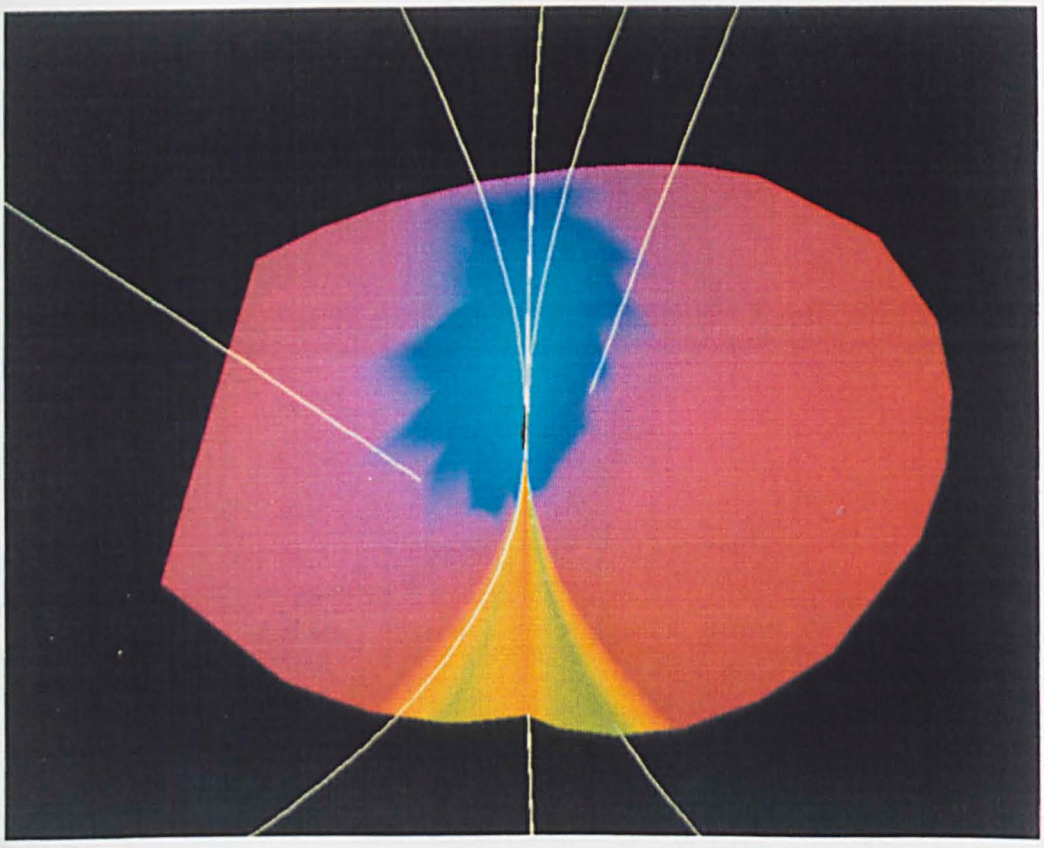


a) both sheets, front view

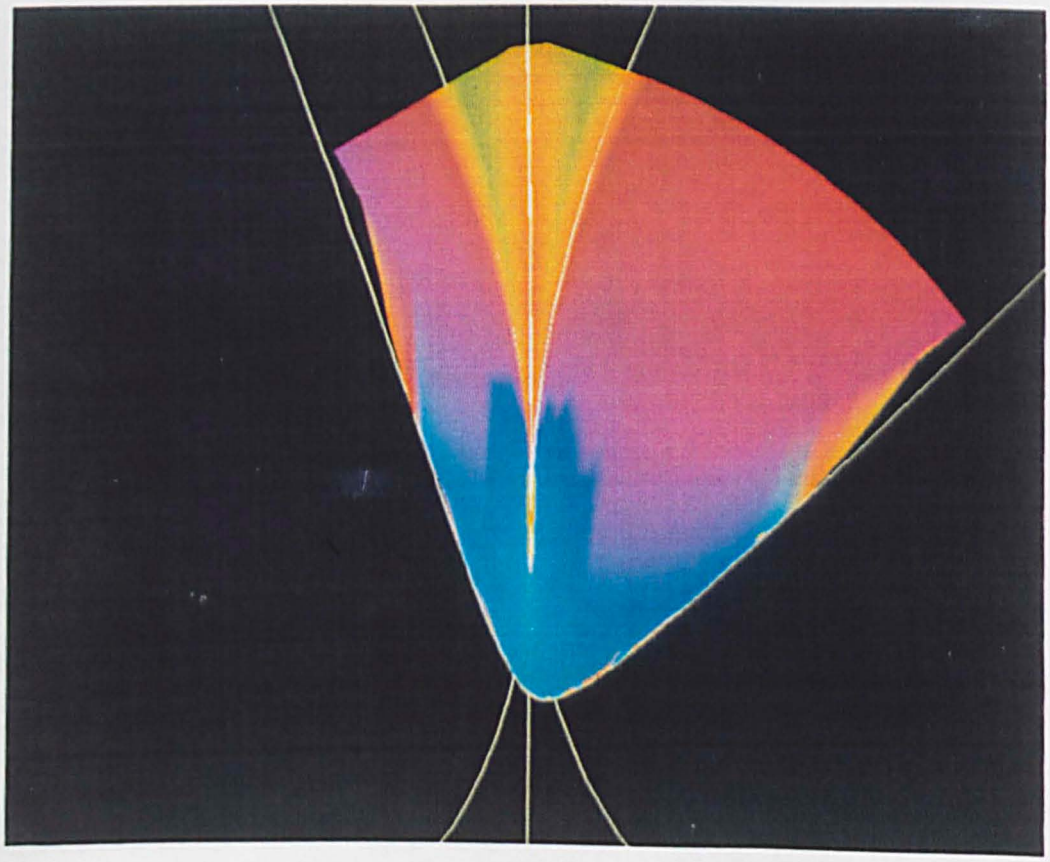


b) both sheets, rear view

fig 8.18) Hyperbolic phase of parabolic umbilic transition

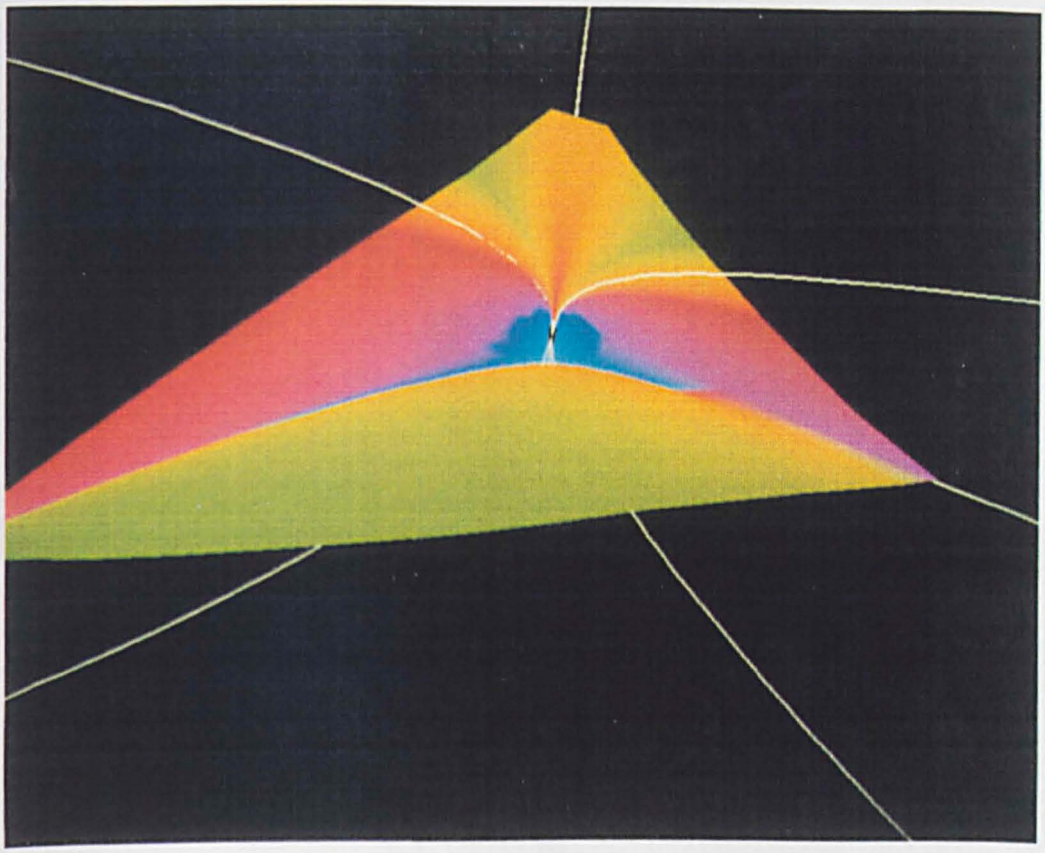


c) sheet 1, front view

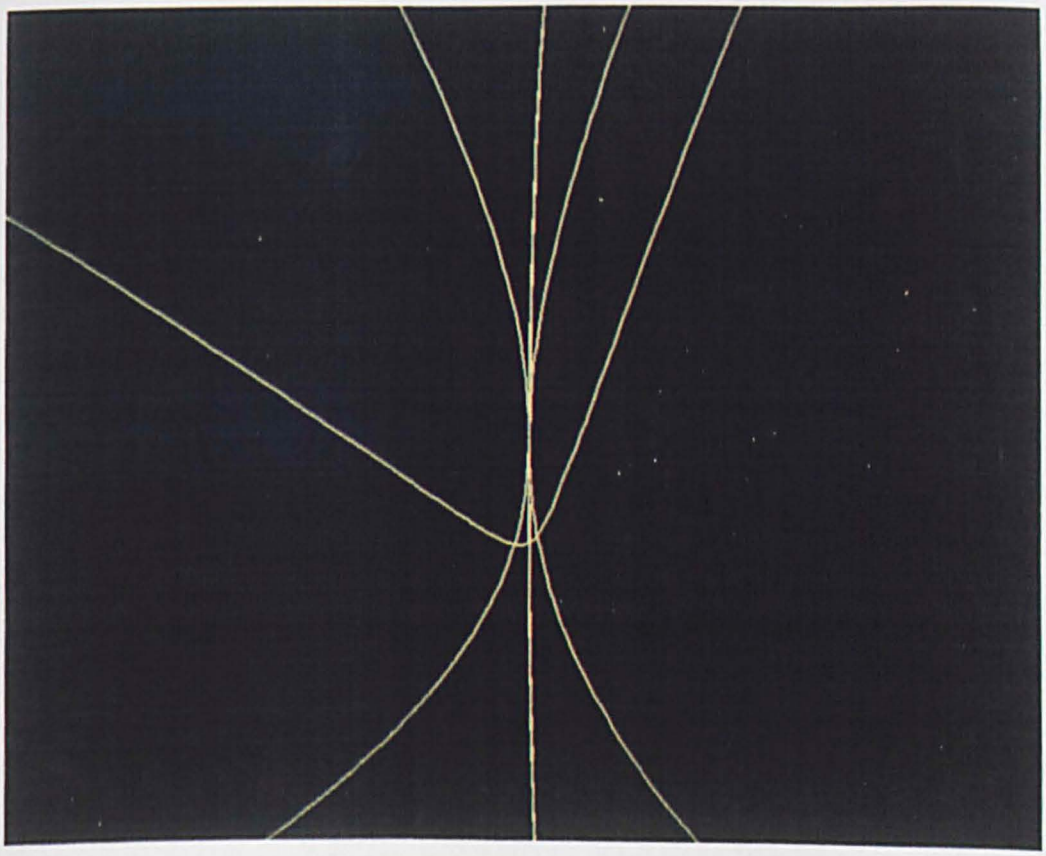


d) sheet 2, rear view

fig 8.18) cont Hyperbolic phase of parabolic umbilic transition



e) sheet 2, view from above



f) the ribs and parabolic lines

fig 8.18) cont Hyperbolic phase of parabolic umbilic transition

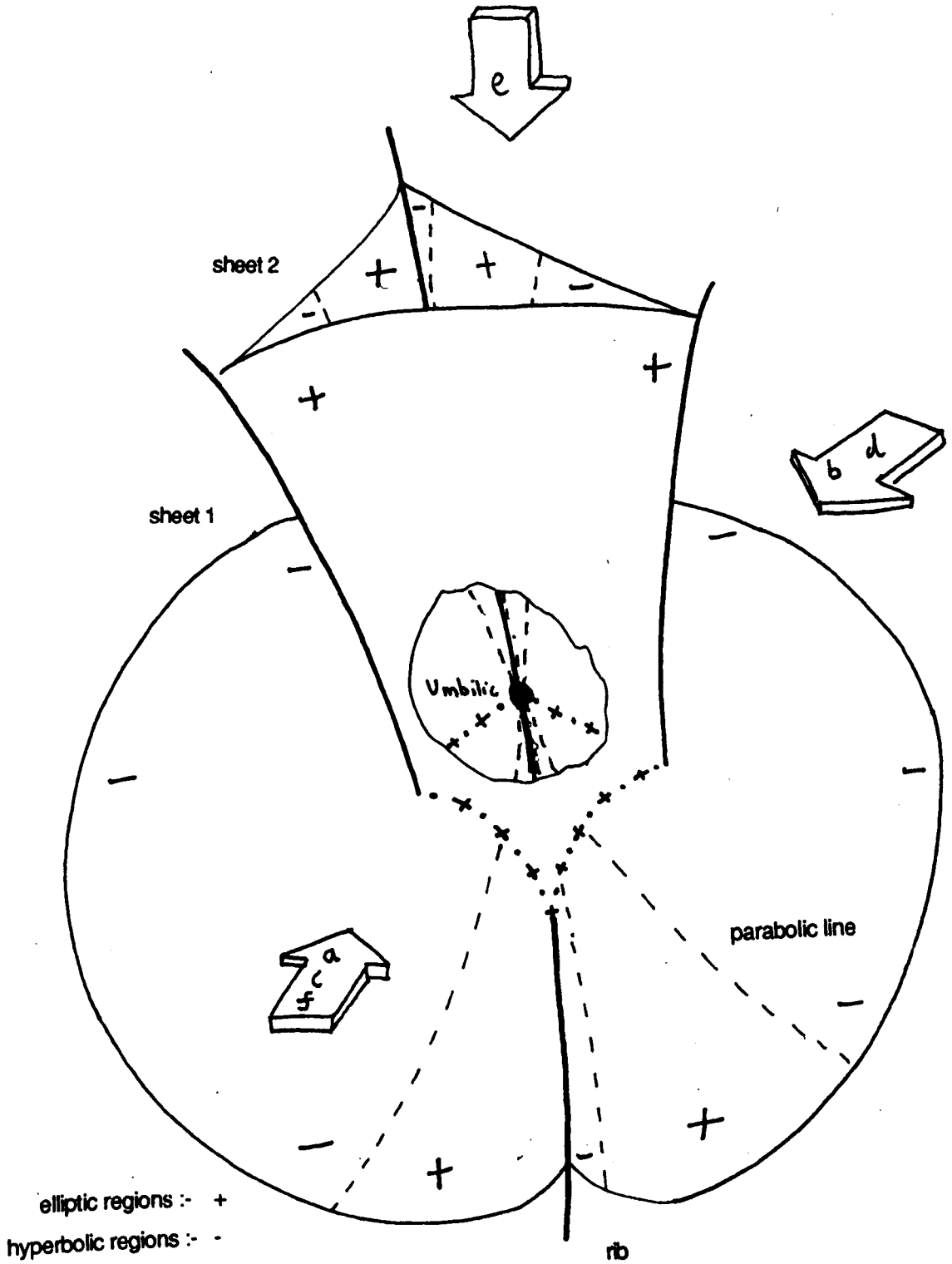


fig 8.18g) Sketch of the umbilic shown in fig 8.18a)-f) indicating the directions of the different views.

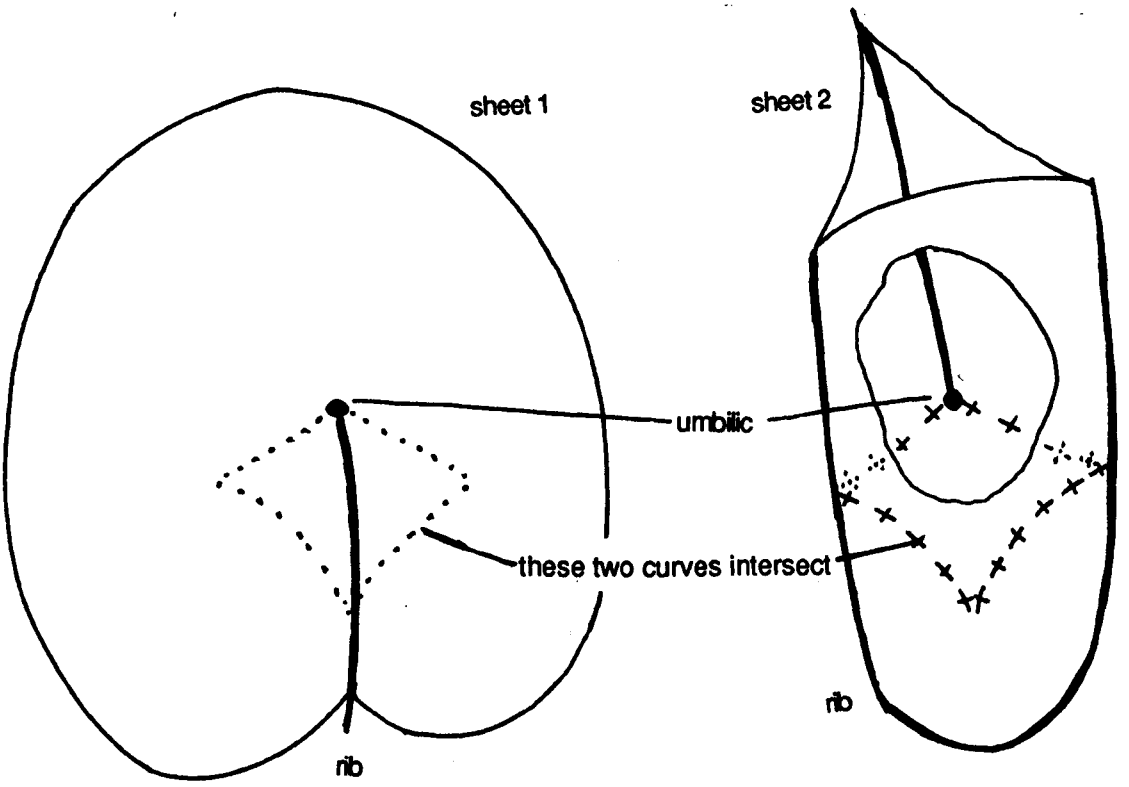


fig 8.19) The two sheets of the focal surface in hyperbolic phase of D_5 transition.

§8.12 Hyperbolic Umbilics

There are three different types of hyperbolic umbilic corresponding to the three different patterns of lines of curvature, "Star", "Monstar" and "Lemon" show in figure 8.2. Two of these, Star and Monstar, have the same number of parabolic lines and it is interesting to find if there are any differences between their focal surfaces.

The principal difference is in their relationship to the surface. We let the sheet 1 correspond to the highest of the two principal curvatures at each point, and hence the lowest radius of curvature. Let the sheet 2 correspond to the lowest of the principal curvatures. Also let π be a plane parallel to the tangent plane of the original surface at the umbilic, chosen such that π passes through the umbilic centre. Figures 8.20a), b) show the relationship between the ribs and parabolic lines of each sheet and the plane π. This sketch was obtained by studying computer generated pictures. For a hyperbolic Star umbilic the rib and parabolic lines of the sheet 1 will lie below π and rib and parabolic lines of the sheet 2 lie above π. In the hyperbolic Monstar case the situation is reversed. This latter case is what holds for a hyperbolic Lemon where there is only one sub-parabolic line. Apart from this

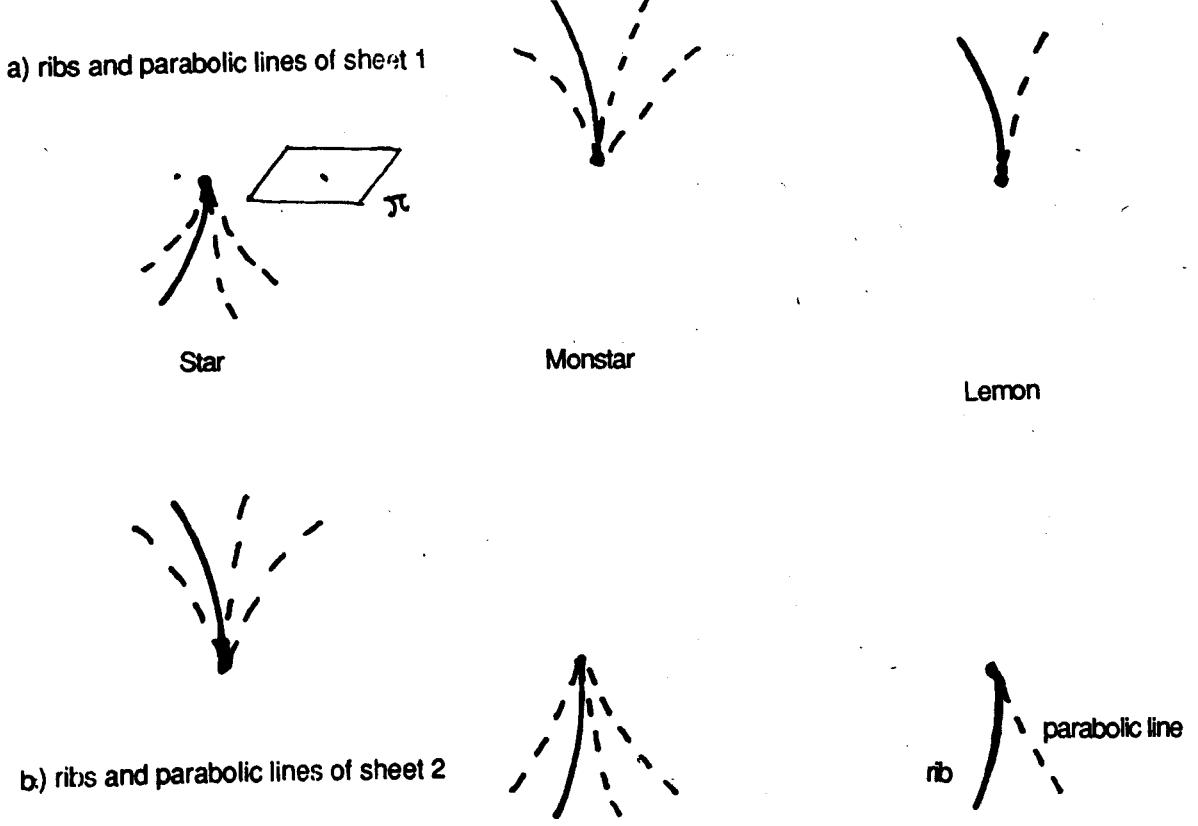
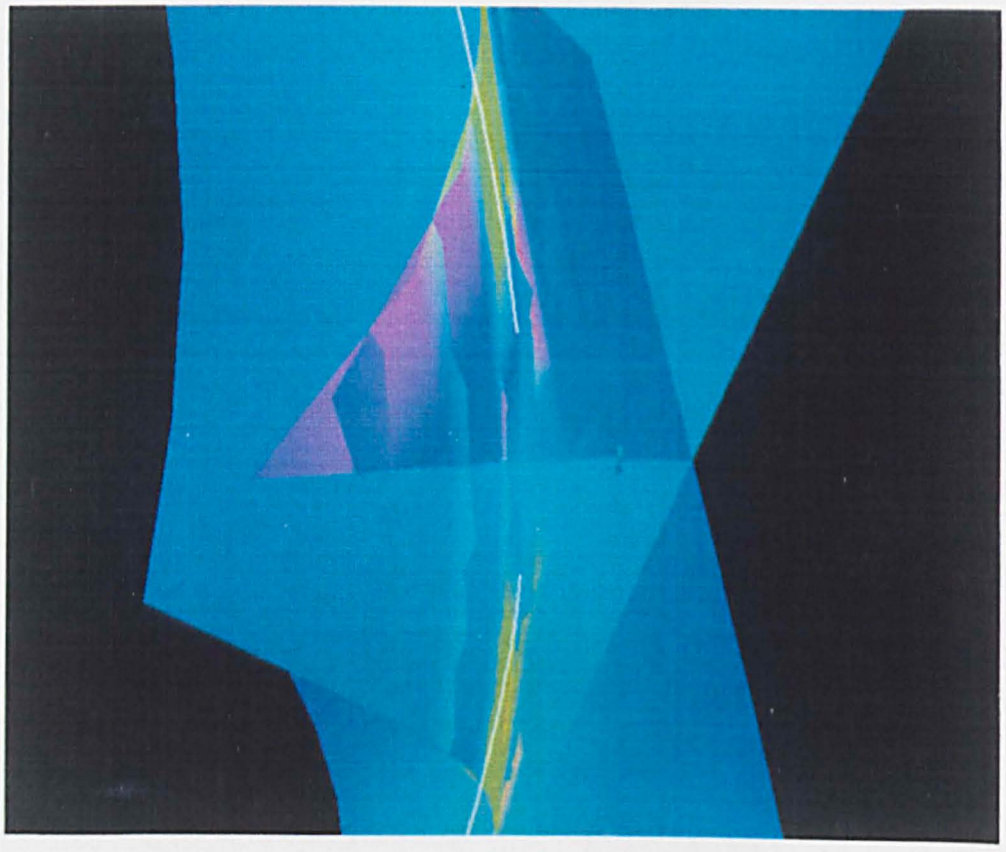


fig 8.20) ribs and parabolic lines for hyperbolic umbilics, assuming original surface lies at bottom of page.

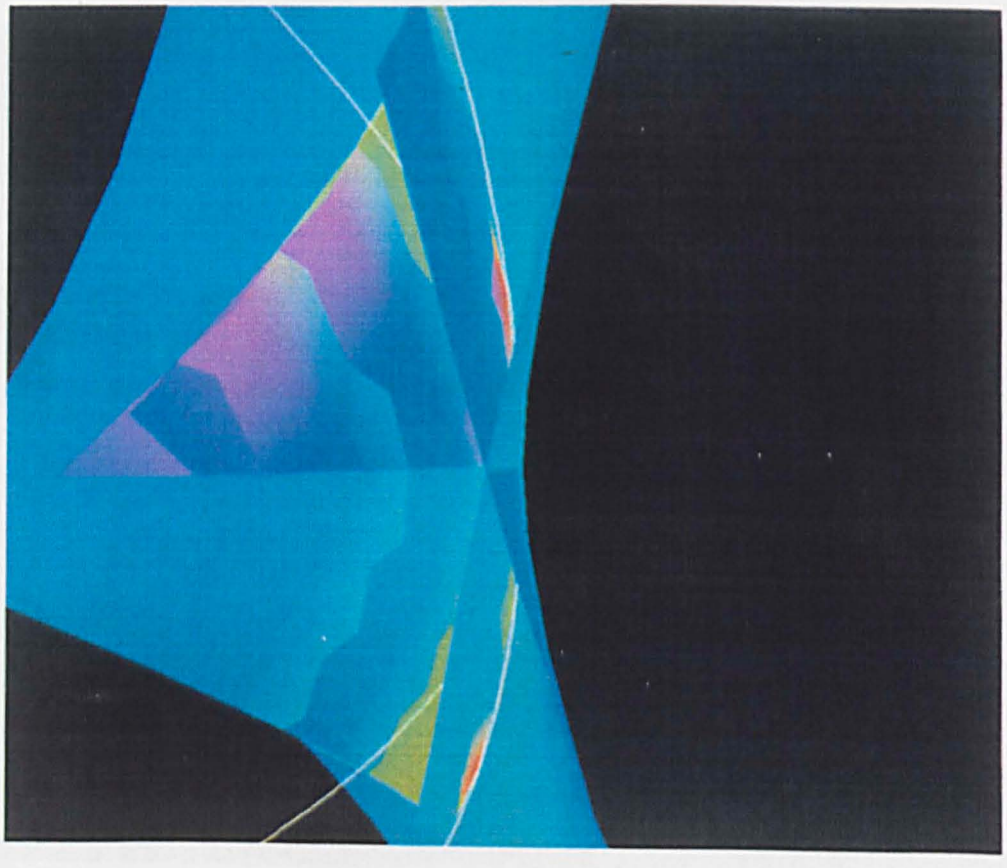
no other essential difference has been observed.

This arrangement ties in with the birth of umbilics transition (§8.6) where a Star and Monstar umbilic annihilate each other. If, say, the hyperbolic Monstar lies above the hyperbolic Star then we define planes π_M, π_S, as above which pass through the respective umbilic centers. The rib of sheet 2 will lie above π_S and below π_M and will shrink down to a point in the transition. The rib of sheet 1 has two parts one above π_M and one below π_S; these will join up to form a smooth rib during the transition.

Just one computer generated example of a hyperbolic umbilic is shown (fig. 8.21)*). This is a Lemon type which has just one rib and one parabolic line (both drawn in white). This picture has been magnified by a factor of ten in the x and y directions. This magnification alters the value but not the sign of the Gaussian curvature. We can clearly see how the two surfaces intersect and also deduce the way the surface curves. This is much easier to see when the surface can be rotated in real time. Note how the rib is curved away from a vertical plane which creates a parabolic line and a region of elliptic points. One sheet of a hyperbolic star umbilic can be seen in figure 8.18c).

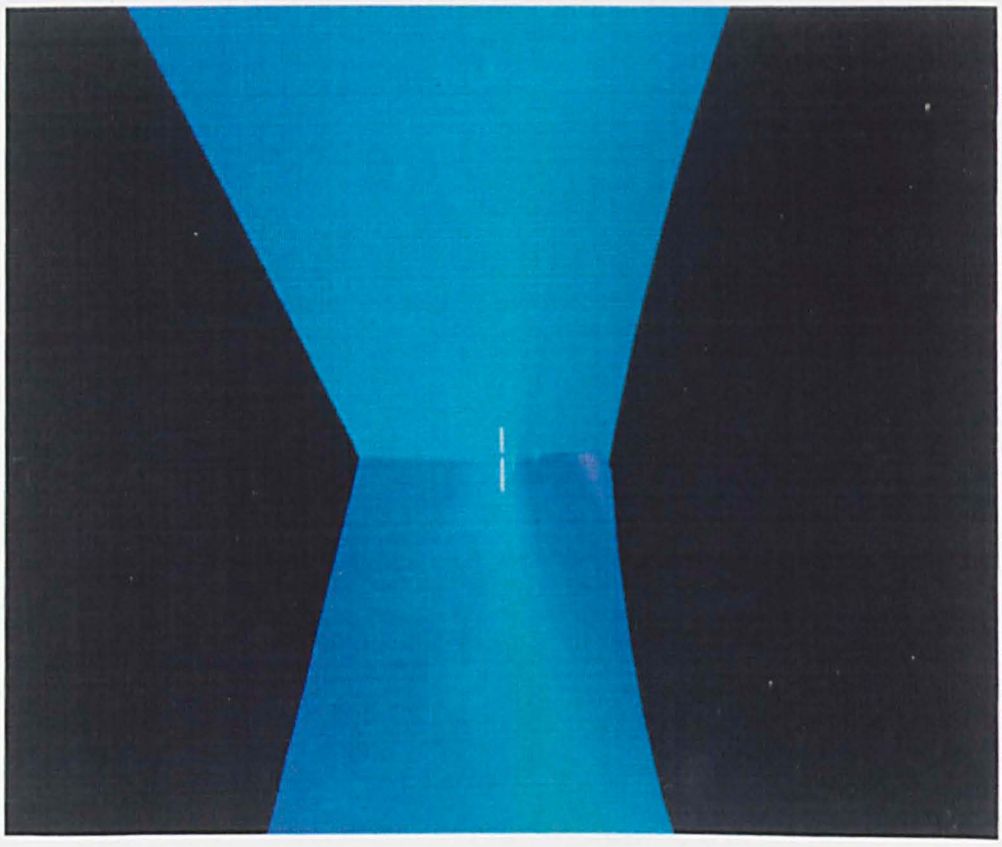


a) both sheets, front view

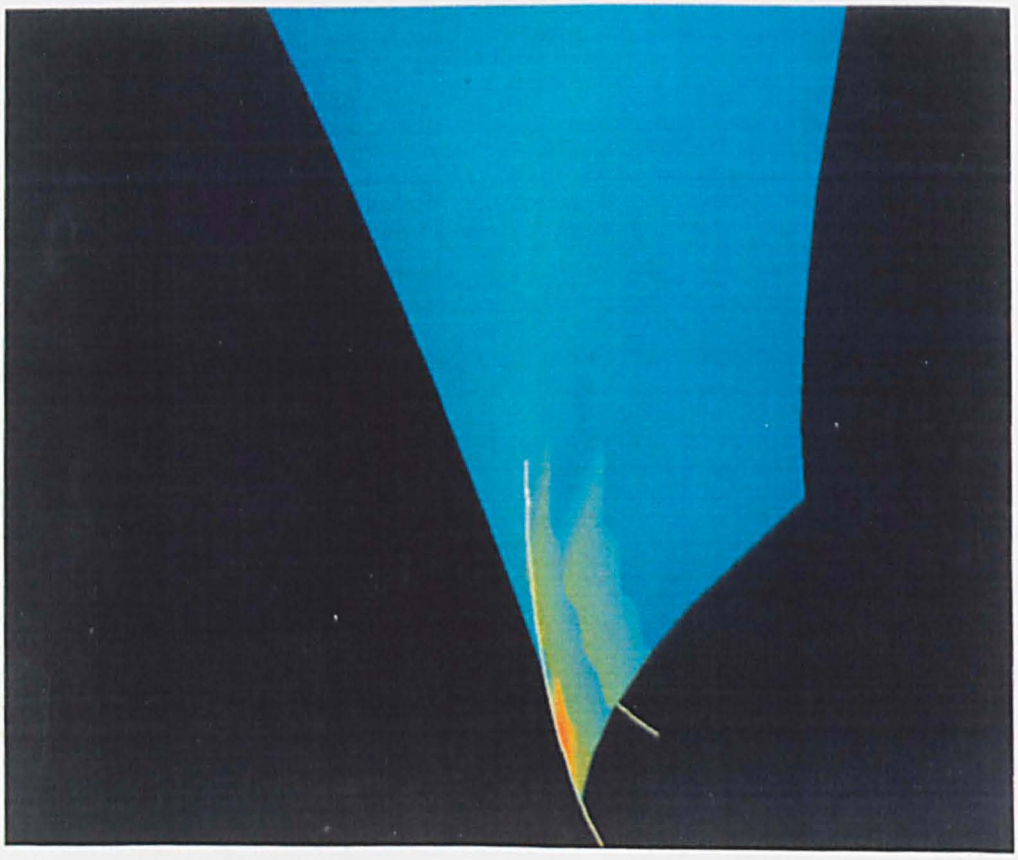


b) both sheets, side-front view

fig 8.21) Hyperbolic Lemon umbilic (magnified in x, y directions)

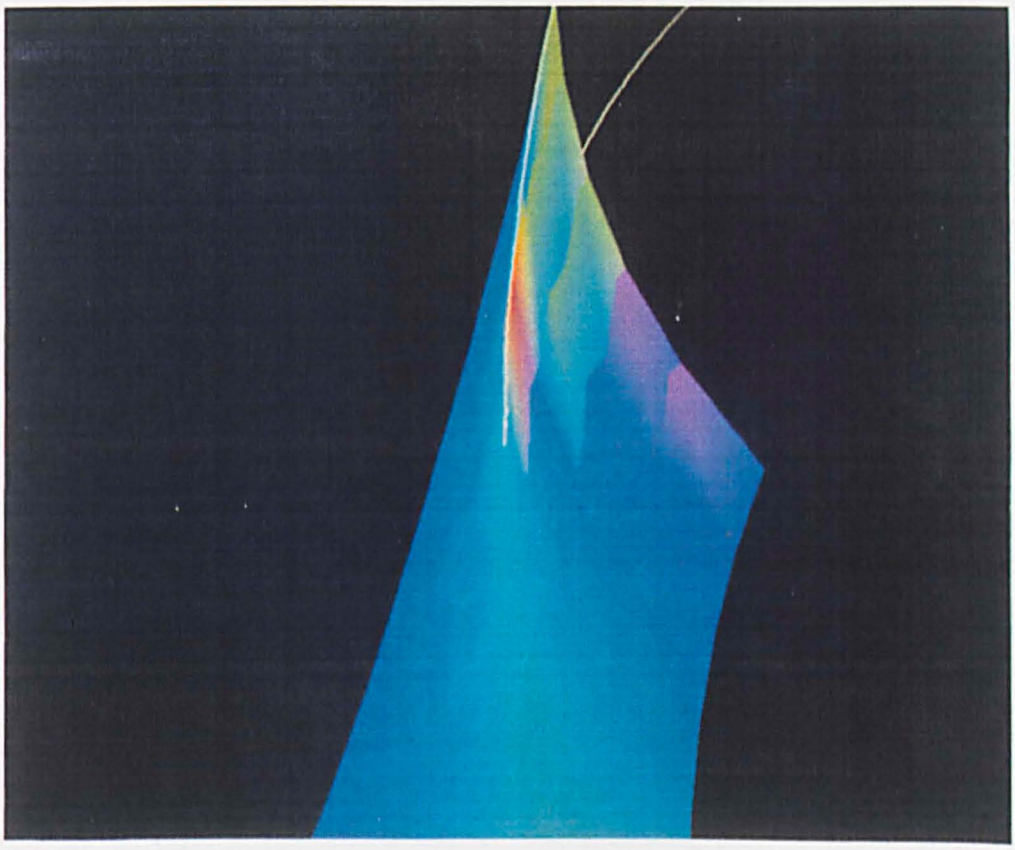


c) both sheets, rear view

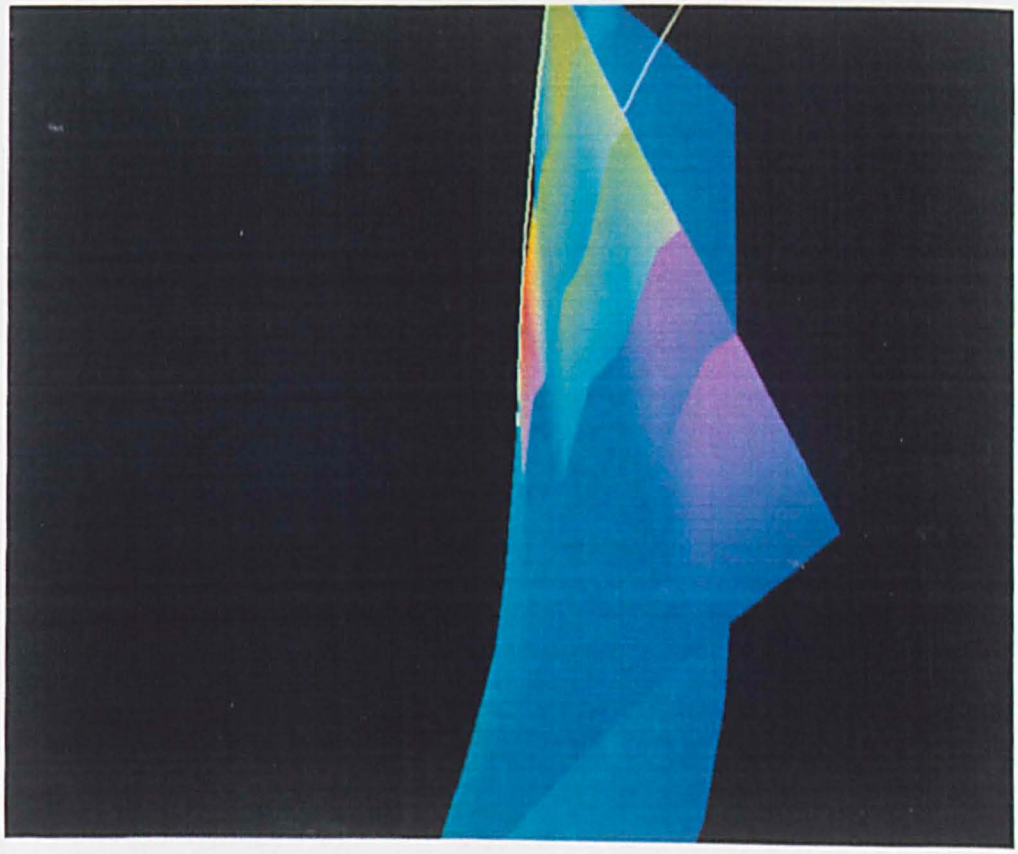


d) sheet 2

fig 8.21) cont Hyperbolic Lemon umbilic (magnified in x, y directions)



e) sheet 1



f) sheet 1

fig 8.21) cont Hyperbolic Lemon umbilic (magnified in x, y directions)

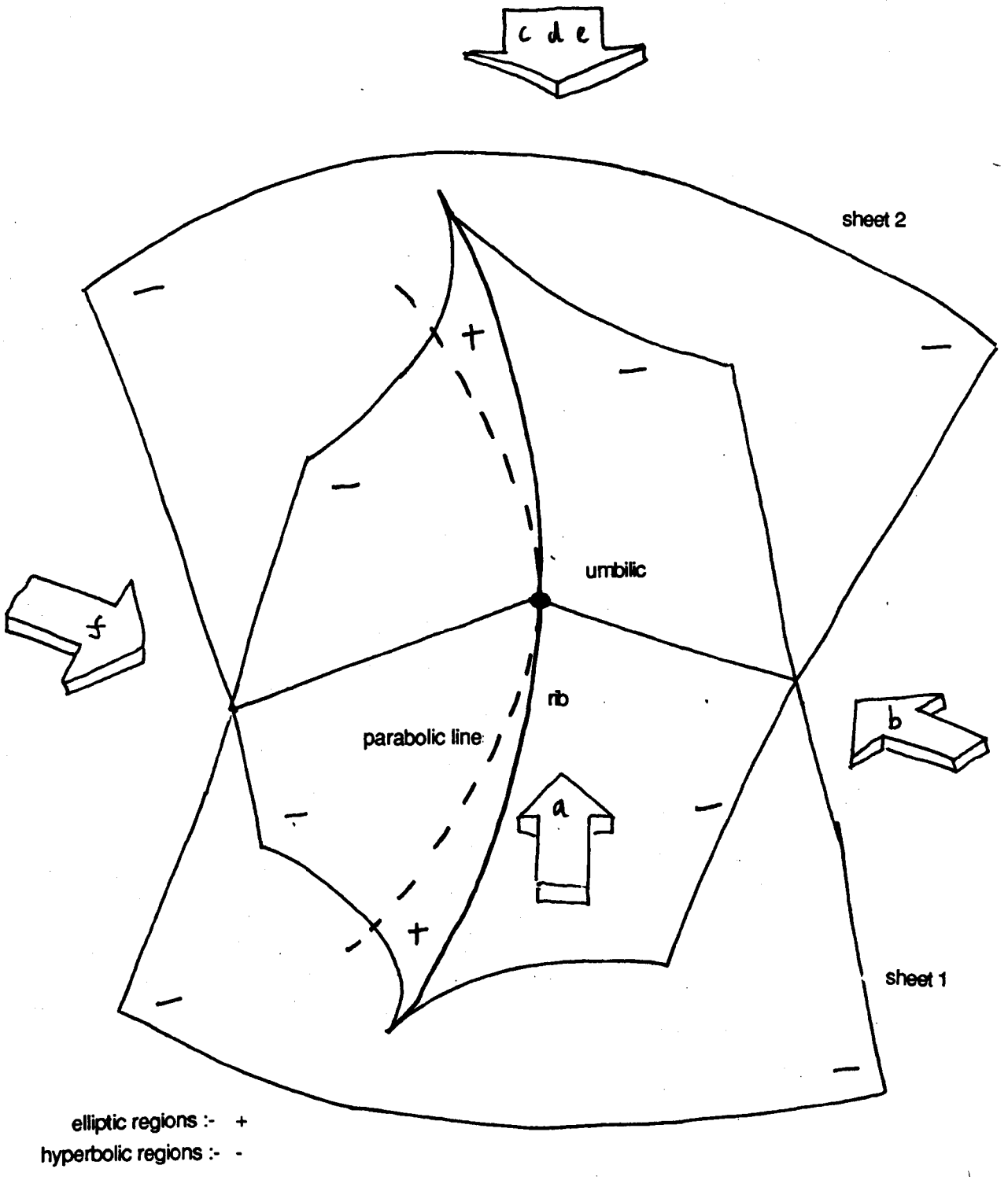


fig 8.2(g) Sketch of the umbilic shown in fig 8.18a)-f) indicating the directions of the different views.

The focal surfaces have proved particularly problematic to draw. As well as the difficulties in drawing the surfaces encountered in chapter 9 there are problems with the geometry of the actual surfaces drawn: non local features often complicate the surface. For instance there may be an other nearby umbilic centre — a slight change in parameters could bring the the two umbilics together in the birth of umbilics transition. Another complication is that the surface might be close to passing through the parabolic umbilic transition so the umbilic might lie inside the pocket shown in figure 8.18)*. The presence of these nearby features prevents too large a patch being used. However if a small patch is used it becomes difficult to distinguish between the two focal sheets or between ribs and parabolic lines. One problem which has occurred in most examples is that one of the ribs and a parabolic line lie close together. The small region of points between the two has not been distinguished by its colour.

There are many other transitions which can be studied:- the orthogonal transition, (the effect of this on a focal surface is subtle as the ribs and parabolic lines of each sheet do not change their arrangement); the birth of umbilics transition and the change between hyperbolic Lemon and hyperbolic Monstar. Other transitions can be studied which do not involve umbilics. For example the when a parabolic line passes through a swallowtail point, the A_5 and the non-generic A_4 transitions.

Chapter 9: The program

A major part of the work has been to write a computer program to draw surfaces and focal surfaces with special points and lines marked on them. The program takes as an input a parametrization of a surface, calculates where the ridges and sub parabolic lines are and then draws them on the surface or corresponding focal surface. For focal surfaces the program also calculates suitable grids of points for representing these complex shapes.

The program can be divided into two main parts: the first part – calculating where the ridges and other features lie in parameter space is performed by a zero following algorithm (§9.2–9.7) similar to that used in the first half of this thesis (Chapters 3 and 4). But it has been modified to cope with the particular problems that occur in this application. The second part is concerned with drawing the surface and the special lines on it (§9.8–9.14).

§9.1 Representing the surface

The main program was written in "C". One of the reasons for this choice of language was that data structures could be used. At each point on the surface a large amount of information is stored in a single unit or "structure". This information includes:

The coordinates in the parameter space;
The principal directions in parameter space defined in two different ways;
The normal to the surface;
The two focal points;
The position of the point on the surface;
structures representing the 1st–4th derivatives of the surface at that point.

Furthermore when calculating features on the focal surface with respect to a given direction the following information was also used:

Two unit length principal directions;
The derivative in the chosen direction of
a) The focal surface,
b) The principal direction field.

This additional information was stored in a separate structure. The use of structures greatly simplifies the program, particularly in the parameters of subroutines.

Instead of having to pass all the information as separate variables just one variable which refers to the whole structure can be used.

Each of the derivatives of a parametrization of a surface, s, at a point is a multi-linear map from a set of vectors $\in \mathbf{R}^2$ to $\mathbf{R}^3$ (§6.1). These derivatives can be represented as a set of real numbers, for example six numbers $(x_s, x_t, y_s, y_t, z_s, z_t)$ are used to represent the first derivative of s at a point, the x, y, z refer to their coordinates in $\mathbf{R}^3$ and the s, t refer to the coordinates in parameter space. The first derivative in the direction $v = (\alpha, \beta) \in \mathbf{R}^2$ is given by

$$ds<v> = (\alpha x_s + \beta x_t,\ \alpha y_s + \beta y_t,\ \alpha z_s + \beta z_t).$$

A single structure is used to hold all six numbers and similar structures are used for the higher derivatives. Together with the position of the point these derivative structures contain all the information required for the calculation. By using sub-routines which evaluate the derivatives on a set of vectors we need never have to think about the underlying Cartesian coordinate system.

The numbers x_s, x_t, ... are just partial derivatives of the coordinate functions so can be easily calculated from the parametrization of the surface. For most surfaces the surface is defined in Monge form and the coefficients of terms up to the fourth order are stored in a file which is read at the beginning of the program.

§9.2 Finding features on the surface

One of the main problems was to find features such as ridges on the surface (or equivalently in parameter space). Many different features need to be found:

a) Ridges corresponding to both focal sheets,
b) parabolic lines on the surface,
c) parabolic lines of both sheets of the focal surface,
d) intersections of each sheet of the focal surface with a sphere centre f_0 radius r_0,
e) lines for which $3d^2V<p^2q> = d^4V<p^4>\ d^2V<q^2>$
(Used for finding A_4 points (lemma 7.2))

Notation 9.1

We will refer to the above objects as **features**. Associated with each feature is a **function** $f : \mathbf{R}^2 \to \mathbf{R}$ for which $f^{-1}(0)$ gives the feature. A number is used in the program to refer to each feature/function. The function is not necessarily globally defined (see §9.4).

The functions corresponding to the above features are

a) $d^3V(x, f(x))<p^3>$,

b) K the Gaussian curvature,

c) $d\underline{P}<q> \cdot \underline{Q}$,

d) $||f(x) - f_0||^2 - r_0^2$,

e) $3d^2V<p^2q> - d^4V<p^4>\ d^2V<q^2>$.

Here p, q are principal directions in parameter space, $\underline{P}$, $\underline{Q}$ are corresponding principal direction on the surface, $f(x)$ is the point on the focal surface and V is the distance squared function. For derivations of these functions see lemmas 7.1, 7.2 and theorem 7.8. These functions have only been given for one sheet of the focal surface, the formulas for the other sheet are similar.

Generically each of the above features form a set of curves in the parameter space. We will only be considering finding features on a bounded patch of the surface. To find these features we use a zero following routine (§9.3). This works by repeatedly being given a point on the feature and then finding another nearby point on it. There are several complications resulting from the presence of umbilics: some of the functions are not globally defined around an umbilic and so make the zero finding more complex (§9.4); most of the features studied either come to an end or split in to two branches at umbilic points. A technique for processing umbilics is discussed in §9.5.

Once the points of the feature have been found in parameter space it is an easy matter to project them onto the surface or focal surface.

§9.3 The zero following routine

The zero following routines in both parts of the the thesis have found each successive point by examining the boundary of a small region for zeros of a function. For the symmetry set (Chapter 4) we follow features in the space $I \times I$ where each coordinate corresponds to a point on the original curve. It seems natural to choose squares with sides parallel to the axis as the regions we examine. When the SS^{-1} is parallel to the axis, (has a turning point), the Symmetry Set has a cusp. Using axis aligned squares aids the detection of these. Here as the coordinates basis has no special significance triangles are used instead. This simplifies the algorithm as only three sides need to be tested instead of four.

The routine will successively work its way along a feature until a boundary or umbilic point is found. Finding the first point on each feature is dealt with in §9.5 below. At the beginning of each step we start with a triangle ABC and a solution

to $f(x) = 0$ on the side AB. The first step is to test to see if there are solutions on the other two sides. There are five possibilities (fig 9.1),

a) Solution on side AC only,
b) solution on side BC only,
c) No solutions,
d) Solutions on both sides,
e) The boundary of the domain has been reached:- stop following.

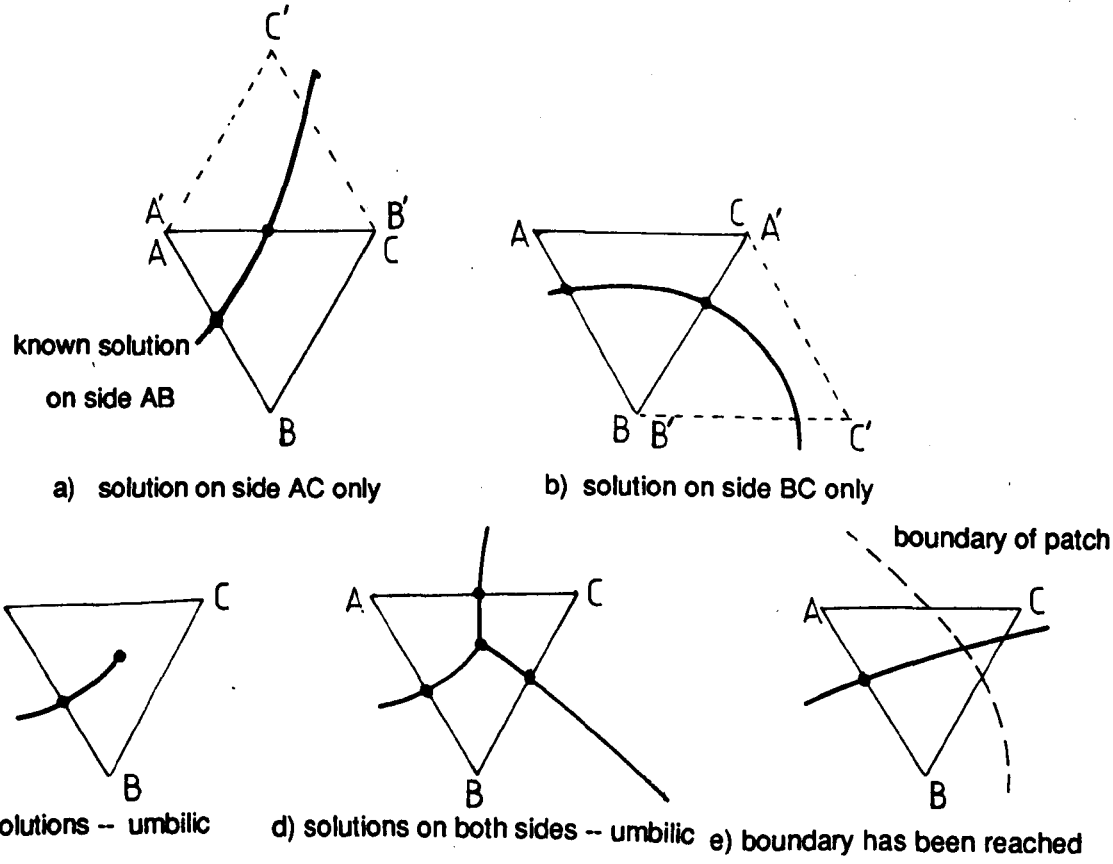


fig. 9.1) Possible cases when looking for solutions round a triangle

The actual positions of the solutions are also found (see §9.4). In the first case a new triangle $A'B'C'$ is created with vertices $A' = A$, $B' = C$ and $C' = A + C - B$. The solution found on the side AC is now on the side $A'B'$ so we can repeat the step starting with triangle $A'B'C'$. Case 2 is similar but here $A' = C$, $B' = B$ and $C' = B + C - A$. For case 5 we simply exit the routine. In cases 3 and 4 we are near an umbilic point. These situations are dealt with in §9.5 below.

§9.4 Finding zeros

A major problem with finding ridges (or sub-parabolic lines) is that the function defining the ridge points depends on the principal directions. But around an umbilic

point we can not define a continuous non-vanishing principal direction vector field. For example if $F : \mathbf{R}^2 \to \mathbf{R}^2$ is a continuous principal direction vector field and if 0 is an umbilic then there is some line $\mathcal{L}$ starting at 0 for which $F(x) = 0$ for all points x on the line. This means the function will have zeros along this line and these do not correspond to the feature of study.

If we take unoriented principal directions around an umbilic there are three main patterns, two of which are shown in (fig. 9.2)*). For computational purposes the third patter "Monstar" is similar to the "Lemon". All three patterns have index $\frac{1}{2}$ or $-\frac{1}{2}$ [Bruce-Fidal] [Porteous-2]. As the indices are non-zero we can not orient the lines to make a continuous non vanishing vector field. The best we can achieve is to have the vector field vanishing or being discontinuous along one line radiating out from the umbilic.

This problem can be overcome by using two different vector fields which vanish along different lines (fig. 9.3)*). At any non umbilical point we can find a neighbourhood of the point where at least one of the vector fields does not vanish. On this region the zeros of the function defined by this vector field correspond exactly to the feature of interest.

In practice two vector fields can be easily found. At each point the principal directions are found by solving an Eigenvalue problem §6.3. This gives two linearly dependent equations for each Eigenvector. For example if $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is a singular matrix, $\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| = 0$, then the equation

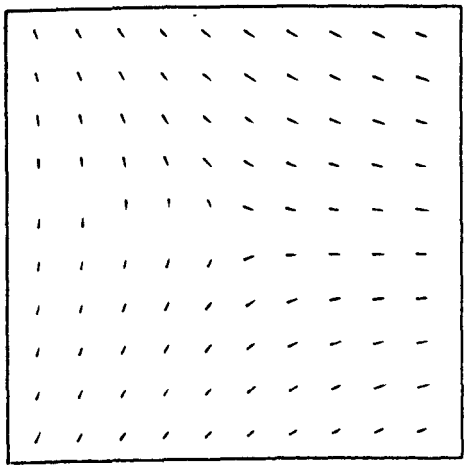
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

has two linearly dependent solutions given by

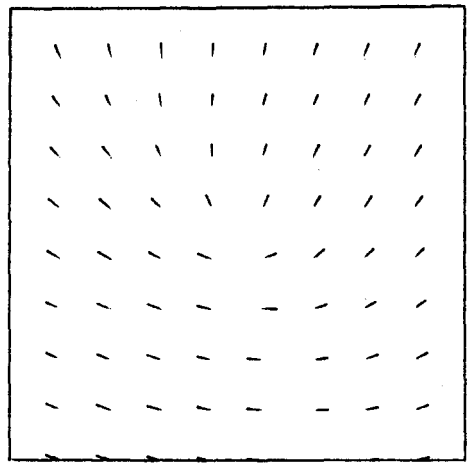
$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ -a \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} d \\ -c \end{pmatrix}.$$

These solutions can be used to form two vector fields. It is possible for one of these to give a degenerate solution, i.e. $x = y = 0$, which will give the points where the vector field vanishes. The only time both give degenerate solutions is when $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is the zero matrix and we have the condition for an umbilic. We can use each of these two solutions to define a vanishing continuous vector field.

When trying to find a zero of a function on a line segment between two points we first have to make sure that the vector field does not vanish on the line segment. First we find the two vectors at the end of the line segment as defined by one of the vector fields. Then we calculate x the cosine of the angle between the two vectors.

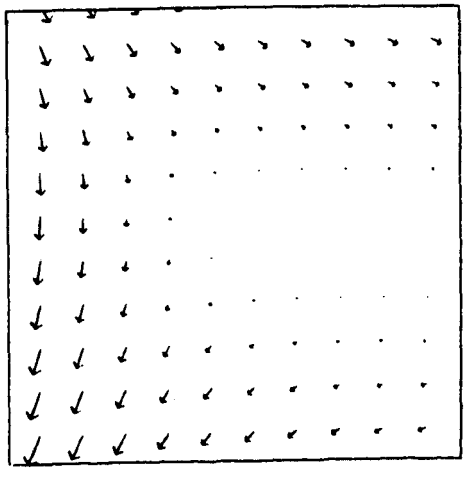


a) Star Pattern



b) Lemon Pattern

Fig 9.2) Unoriented principal directions around an umbilic



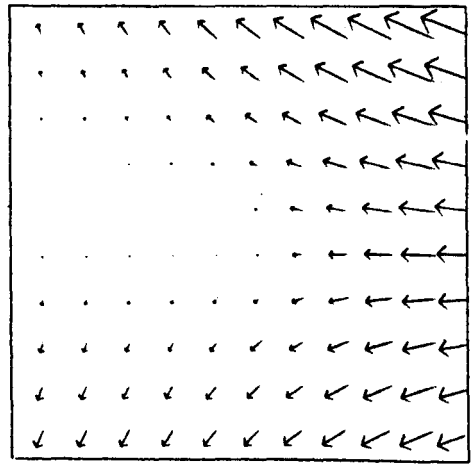
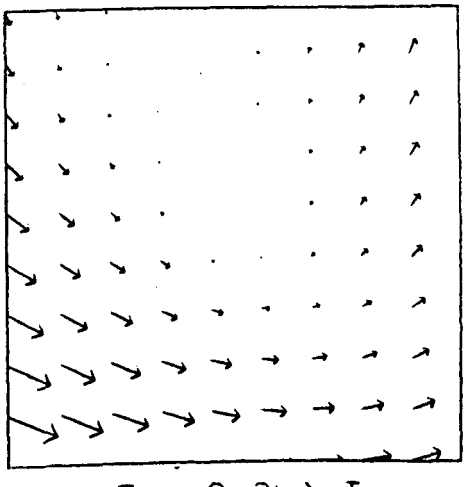


Fig 9.3a) Two orientated principal direction vectors field for the Star pattern



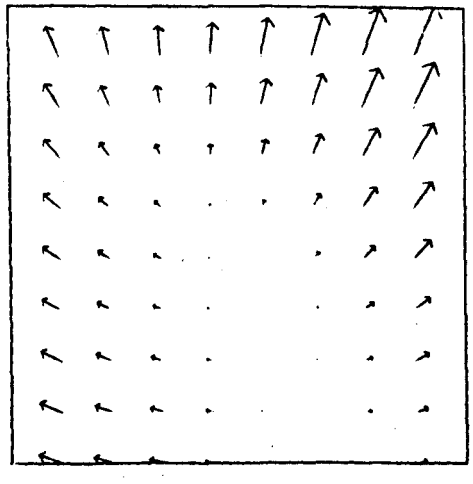


Fig 9.3b) Two orientated principal direction vectors field for the Lemon pattern

If x is greater than some fixed positive number (which we shall call dot_tol which was given the value 0.9) then we assume that the vector field is non-vanishing between these two points (fig 9.4). If x is less than $-dot_tol$ then the vector field must have turned through nearly 180° degrees hence the vector field has probably vanished between the two points (fig. 9.5a). In this situation we try the other vector field to see if better results can be obtained (fig. 9.5b). If x lies between $-dot_tol$ and dot_tol (fig. 9.6a) we can not be sure if the vector field has vanished or not so we divide the line segment in half an check each part separately (fig 9.6b). This process is then repeated recursively until we have non-vanishing vector fields on each line segment. Now we can use this vector field to define our function and look for its zeros. If the function does not depend on the vector fields, for example the Gaussian curvature, then we can skip the previous step.

fig. 9.4) Small change in direction of vector field along an interval

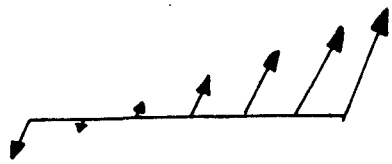


fig. 9.5a) Large change in direction of vector field along an interval

fig. 9.5b) the other vector field in this case

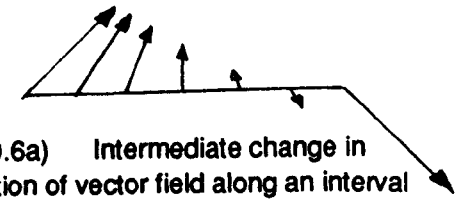


fig. 9.6a) Intermediate change in direction of vector field along an interval

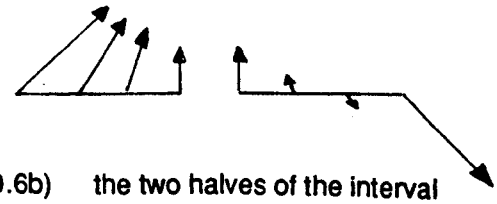


fig. 9.6b) the two halves of the interval

The zeros are found by looking for a change of sign. If there is one a simple subdivision algorithm is used. This divides the interval in half and we check each half for a zero. This process is repeated a few times to converge to the position of the zero. This method works quickly and without problem so this simple algorithm was used instead of a more complicated method such as the Newton's method.

§9.5 Finding starting points

Before using the zero following routine some starting points must be found. As we are mainly concerned with surface patches it is sensible to see where the feature

crosses the boundary of the patch. This is done by finding a parametrization of the boundary and moving round it looking for points on the feature.

Just finding points on the boundary will not always find all the required starting points. We can improve on this by examining the geometry of the ridges. In section 8.6 we see that umbilic points are born in pairs both being of hyperbolic type. Between the two umbilics there will be a new isolated section of ridge points (fig. 9.7a). An other special case is where a small isolated loop of ridge points is created during a non-generic A_4 transition (fig. 9.7b) [Porteous–2]. There will be two A_4 points on the loop. These can be found by following the curve defined by $d^4V<p^4>d^2V<q^2> = 3(d^3V<p^2q>)^2$ and finding points where $d^3V<p^3> = 0$ on the curve (see §9.7).

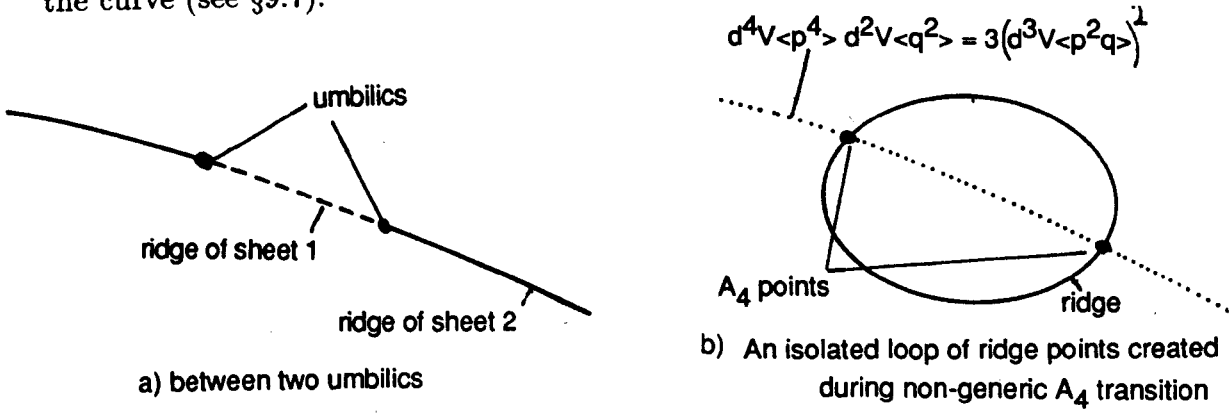


fig. 9.7) Isolated segments of ridge points

Lists are kept of the umbilic and A_4 points as they are found. We can then use the points in these lists as new starting points for the following routine. Even using these points there is still the global problem of whether all the ridges and sub-parabolic lines will be found but as we have just been looking at the local structure this global problem has not arisen.

§9.6 Processing Umbilics

The first version of the program was not very clever when it came to umbilics. An umbilic was simply regarded as a triangle where there were a odd number of solutions round its boundary. For some examples, such as parabolic umbilic where two solutions were close together, the program would not find all the solutions.

An improved version where the actual position of the umbilic was found was later implemented. When an odd number of solutions were found round a triangle the umbilic processing routine was run. This first finds a good approximation to the position of the umbilic. If κ_1 is the higher of the two principal curvatures and κ_2 the lower, then the function $\delta\kappa = \kappa_1 - \kappa_2$ is a smooth function which is zero at

umbilics and positive everywhere else. A simple algorithm for finding the minimum of this function given a point (x_0, y_0) near the umbilic works as follows. A small number ϵ is chosen and the points $(x_0, y_0 \pm \epsilon)$, $(x_0 \pm \epsilon, y_0)$ are tested to see which has the lowest value of $\delta\kappa$. The point with the lowest value then becomes the new starting point (x_1, y_1) and the number ϵ is halved. This process is repeated a small number of times and provided $\delta\kappa$ is well behaved we should converge quickly to the umbilic.

Once the position of the umbilic is found we find the directions of the various features passing through the umbilic. To find these directions a small circle round the umbilic is used. This circle is divided into a number of segments and each segment is examined for solutions x_i of the functions defining the ridges/sub-parabolic lines of both sheets. The information about these solutions is stored in a data structure about the umbilic. In the actual program 48 segments were found to be sufficient.

Each of the solutions x_i can form the starting or ending point for pieces of the features and flags are used to see which solutions have been used in the following routine. Normally the umbilic processing routine is called from the following routine and the last point D found is passed as an argument. For each solution x_i the angle ϕ_i round the umbilic of the point D is found(fig. 9.8). The flag of the solution with the minimum angle is set to "solution used."This angle should be quite small

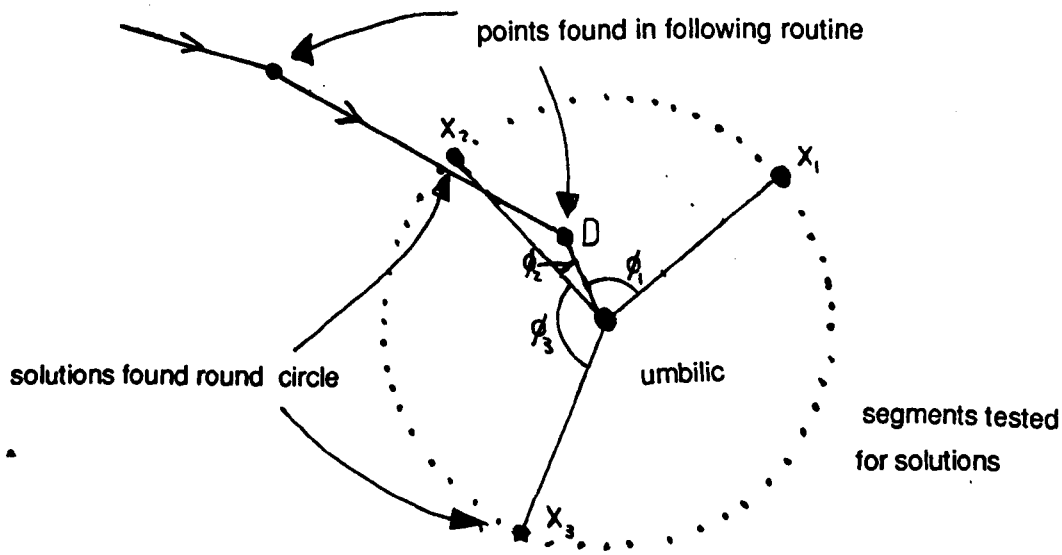


fig. 9.8) The correspondance between solutions found round a circle and the points found in following routine

as features passing through an umbilic can normally be approximated by a straight line. Once this solution has been checked we can see which of the other solutions

are yet to be used and these can then be used as starting points for the following routine.

§9.7 Finding where two features cross

In several places it is useful to find points where two different features cross, i.e. a point which is a solution to $f_1(x) = 0$, $f_2(x) = 0$ for two function $f_1, f_2 : \mathbf{R}^2 \to \mathbf{R}$. For instance at an A_4 point we require that both $f_1 = d^3V<p^3>$ and $f_2 = 3(d^3V<p^2q>)^2 - d^4V<p^4>\ d^2V<q^2>$ are zero (lemma 7.2).

In a generic situation the two features will cross transversely (fig 9.9). A

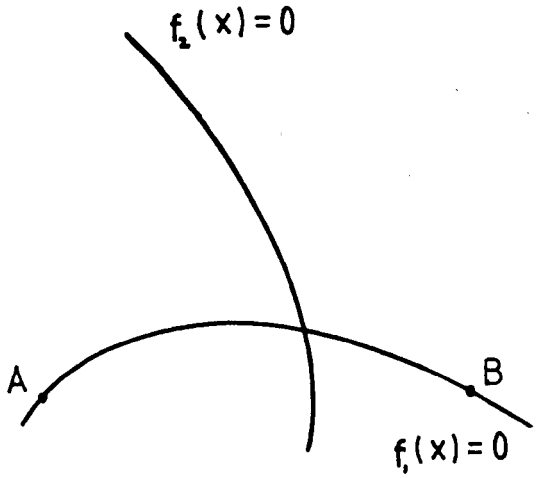


fig. 9.9) Two features crossing transversely

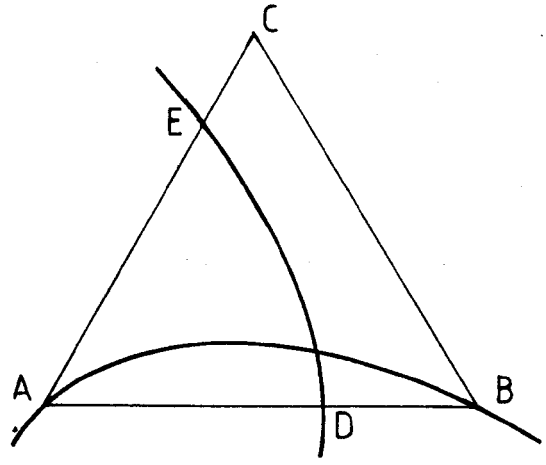


fig. 9.10) Solutions D, E, to $f_1(x) = 0$ around a triangle ABC

routine for finding such crossing points is outlined below. The routine is started by being given two functions f_1, f_2 and two solutions A, B to $f_1(x) = 0$. First the routine attempts to find a solution D for $f_2(x) = 0$ on the line AB. If no solution is found the routine ends immediately. Otherwise the next step is to find a second solution to $f_2(x) = 0$. For this we construct an equilateral triangle ABC where $C = \frac{1}{2}(A + B) + \frac{\sqrt{3}}{2}V$, and V is the vector $B - A$ rotated round by 90° (fig 9.10). Now we find the other solution E to $f_2(x) = 0$ which must lie on either AC or BC. If f_1 has a zero between D and E then there will be a crossing point in the triangle. If no crossing point is found then we try a different triangle with $C = \frac{1}{2}(A + B) - \frac{\sqrt{3}}{2}V$ (fig 9.11). If the first feature is fairly straight or the interval AB is small the crossing point will be contained within one of these two triangles. But if neither returns a solution we try successively larger triangles with $C = \frac{1}{2}(A + B) + m\frac{\sqrt{3}}{2}V, m = 2, -2, 3, -3, \ldots, n, -n$.

When the points D, E have been found we call the routine recursively with $A' = E$, $B' = D$, $f_1' = f_2$ and $f_2' = f_1$ (fig. 9.12). The recursion continues until a

fixed limit is reached and when this happens the point D is returned as the final solution. It was found that not many levels of recursion were needed. Two or three were sufficient to give good results.

An alternative method would be to follow one of the features with a much smaller step size and look for a change in sign in the other function. This might prove to be slower as a point needs to be found for each step. In the above routine two points need to be found for each level of recursion but the triangles decrease rapidly in size. Therefore less points need to be found for the required accuracy. One reason for choosing this method was its symmetry. First an approximate solution is found lying on one feature and then and then an approximate solution is found

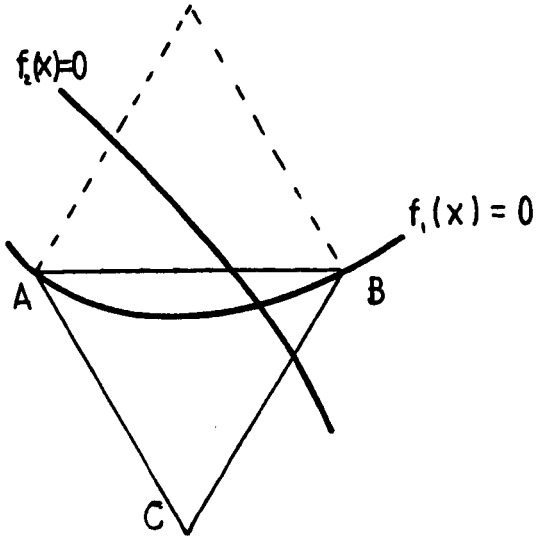


fig. 9.11) Trying a different triangle if no crossing point found

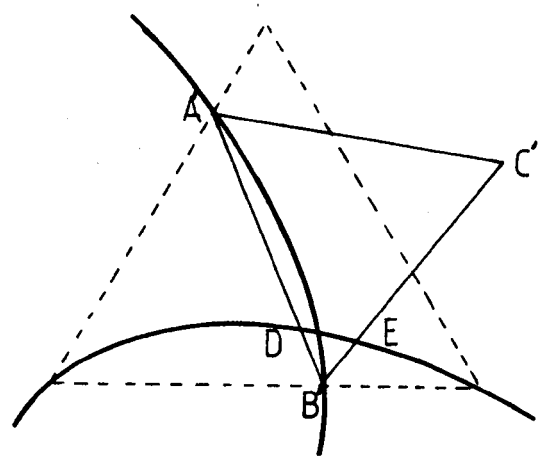


fig. 9.12) The next level of recursion after fig. 9.10)

lying on the other feature. This prevents bias towards one of the features.

§9.8 Introduction to drawing surfaces

Before we start to discuss the methods for drawing surfaces we must know the types of surfaces to be viewed. Here they are all small patches of parametrized surfaces, lying in three dimensional space. They are either smooth surfaces or their associated focal surfaces. We will write the parametrization as $f : \mathbf{R}^2 \rightarrow \mathbf{R}^3$. The focal surfaces contain many singularities such as cuspidal edges, umbilics and A_4 points. These may cause problems in drawing. The ability to move around the surface in real time is essential to understanding the geometry of these singularities especially near umbilics. Of particular interest are the cuspidal edges (ribs) and parabolic lines, together with their relationship to the Gaussian curvature. We can easily calculate the tangents and Gaussian curvature at any point (see chapter 7),

but the calculations for higher derivatives become rather complex. Naturally we can only carry out these calculations on a finite number of points. As there is little intuitive feel for the effect the parameters defining the surface have on the resultant image, a large number of examples with slightly different parameters need to be viewed. Therefore it is advantageous to have as short a time as possible from when the parameters are specified to when the image can be viewed. Also it is useful to know the types of hardware and software that are available for viewing 3D surfaces. Two different computer systems were available for viewing surfaces:-

The *CATIA* interactive computer aided design package running on an IBM5080 graphics workstation [Catia].

An *Iris 4D GT* graphics workstation. Programs written in *C* can access the *GT* graphics library to display objects [Iris].

The *Iris* only arrived towards the end of the project so most of the programs were written with the *CATIA* system in mind. These were then slightly modified to suit the *Iris*.

§9.8.1 The CATIA system

As finding points on a focal surface requires complex calculations the focal surfaces can not be created interactively. The data needs to be created by an external program. In *CATIA* library routines are available which enable a *FORTRAN* program to create some of the objects. Those that can be created include:-

1) Points;

2) Lines;

3) Planes;

4) *Polynomial curves*: curves defined by a one variable polynomial of a fixed degree in either $\mathbf{R}^2$ or $\mathbf{R}^3$;

5) *B-spline curves*: given a set of points a B-spline polynomial curve can be created which passes close to each point. This curve does not actually pass through all the data points;

6) A *Net* of points: an $m \times n$ grid of points is defined. At each point tangents and principal curvatures can be specified;

7) *Polynomial surfaces*: defined by a two variable polynomial in $\mathbf{R}^3$,

8) *B-spline surfaces*: given an $m \times n$ set of points a polynomial surface is fitted. As for curves this does not actually pass through all the points;

Options 4), 5), 7) and 8) were not available until halfway through the project. Also the information we have about the surface does not lend itself to these types of representation. Once the above objects have been created they can be operated on interactively. The following options are available:-

9) Planes can be restricted by a set of lines or curves to form a *face*. Only the parts of the plane contained inside the lines will be drawn;

10) Similarly *surfaces* can be restricted by curves or lines defined in their parameter space;

11) *Nets* can be transformed into *surfaces*.

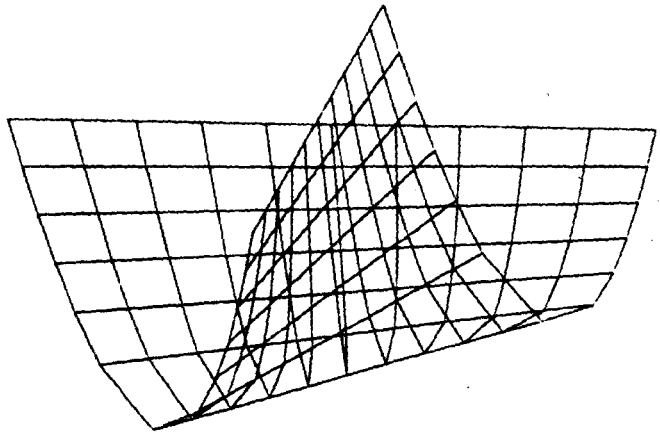
12) *Surfaces* can be transformed into *Polygonal surfaces* which are made up of a large number of polygons.

There are many other interactive options. As mentioned above, to decrease the time taken for the creation of each object it is desirable to reduce the number of interactive options used to a minimum. To help view objects they can be drawn in one of eight colours and can be shifted, rotated and enlarged in real time. *Nets* are drawn as a set of lines connecting adjacent points (fig 9.13a). *Surfaces* are drawn by representational isoparametric curves (fig 9.13b). The hidden lines can be removed from *polygonal surfaces* at the expense of losing the ability to rotate in real time (fig 9.13c). It is also possible to shade or ray-trace *polygonal surfaces* but these take a considerable time to produce the pictures; also the shading algorithm could not cope well with the self-intersections of focal surfaces. For each image several different combinations of lighting need to be tried to produce a reasonable image.

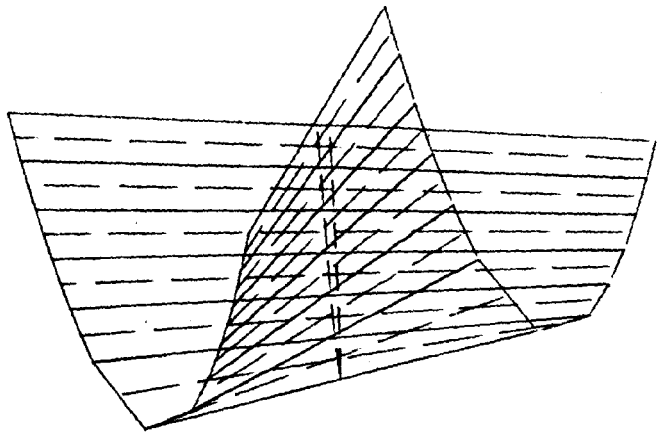
One major restriction is that all the *nets* and *surfaces* are based on rectangular grids of points. These do not fit well with some of the regions we are interested in, which may, for example, be triangular. It is possible to restrict a square to a triangle but this always requires interactive operations and the patch may be singular along the edges, which may cause computational problems.

§9.8.2 The Iris system

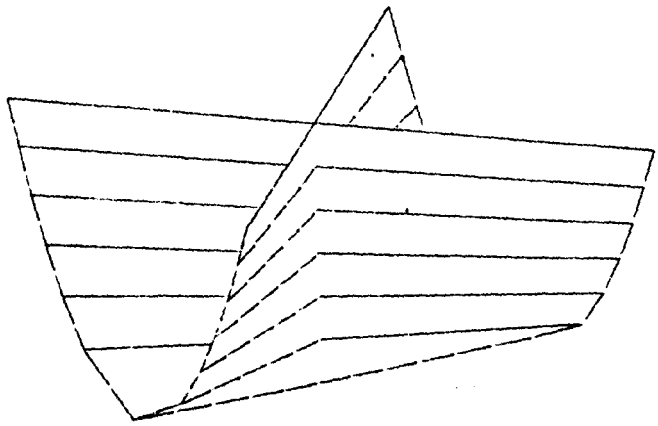
On the *Iris* curves consist of lines joining a set of points (fig 9.14a). Surfaces can be defined by a sequence of points with a triangular face drawn in between each 3 successive points (fig 9.14b). Such a surface is called a *tmesh*. At each



a) draw as a *Net*



b) draw as a *Surface*



c) draw as a *pologonal surface*

fig. 9.13)* Different representations of a crosscap surface in *CATIA*

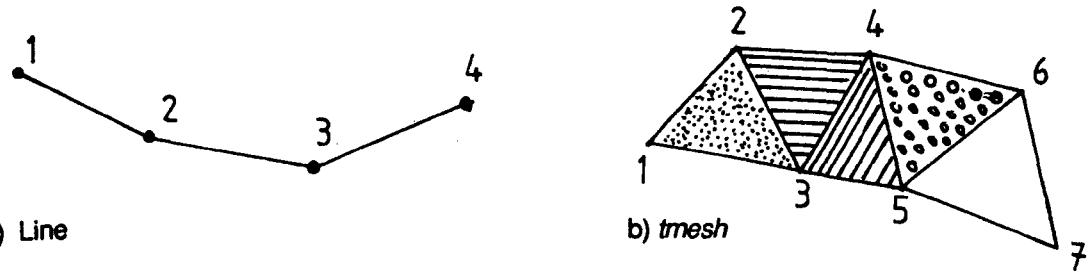


fig. 9.14) Lines and *tmeshs* on the *Iris*

data point a colour can be specified by 3 eight bit RGB values. The colours of points inside a face are determined by interpolation. It is also possible to specify the material charcterestics of the object and the position of light sources and then obtain a shaded picture.

This system has many advantages over *CATIA*:-

Use of a graphics library from a program written in *C* gives a lot of flexibility in the format of the data, as the viewing program can be written to suit the data.

The *CATIA* system is primarily an interactive system. This means that the same sequence of operations needs to be carried out to create each model viewed. In a *C* program these operations can be carried out automatically, thus saving a lot of time.

Z-buffering is implemented in hardware which enables near instantaneous rendering of a 3D scene. This allows shaded or solid colour objects to be transformed in real time.

§9.9 Drawing the image of a curve in parameter space

This can be easily accomplished in both systems by calculating successive points on the curve and then drawing the lines connecting these points in the image. On the *CATIA* system ribs were coloured red and parabolic lines were coloured yellow.

§9.10 Rectangular grids

The simplest method of drawing a surface is to find a rectangular grid of points in parameter space and then draw the grid lines on the surface. Figure 9.15)* illustrates three such grids. The lowest is a parametrized surface and the other two are the corresponding focal surfaces.

A simple improvement can be made by cutting the grid lines whenever they cross a ridge or sub-parabolic line (see §9.7). This is carried out in the parameter

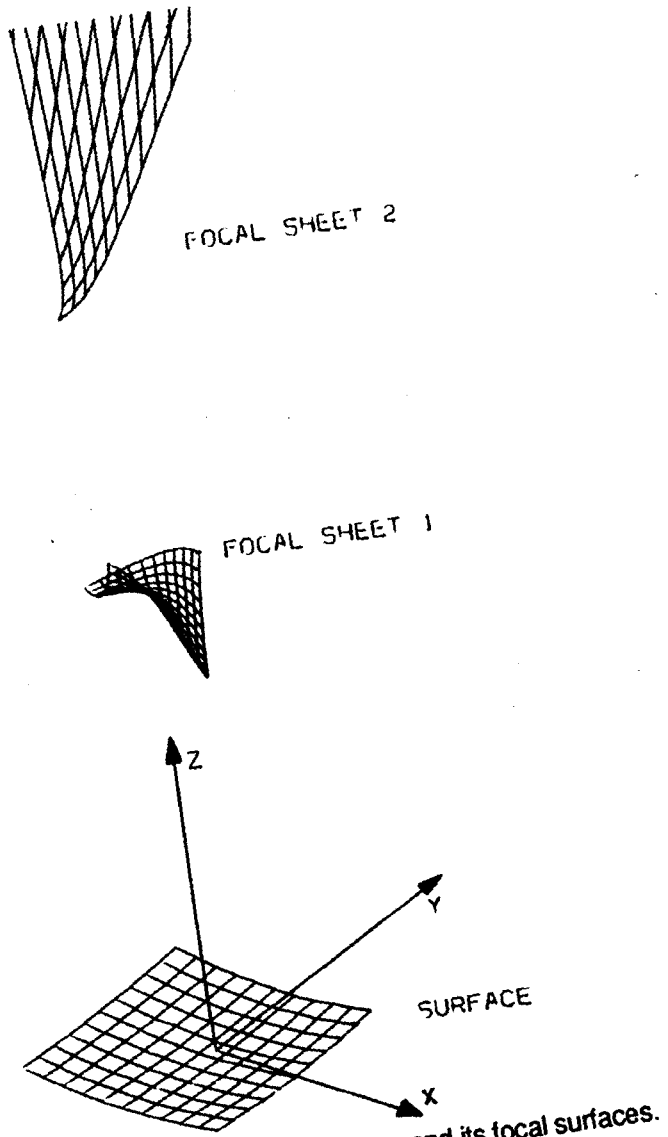


fig. 9.15)* A rectangular grid on a surface and its focal surfaces.

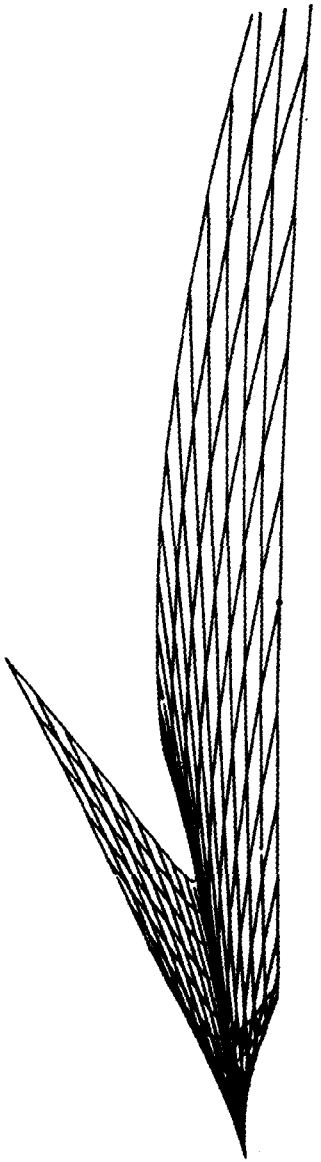


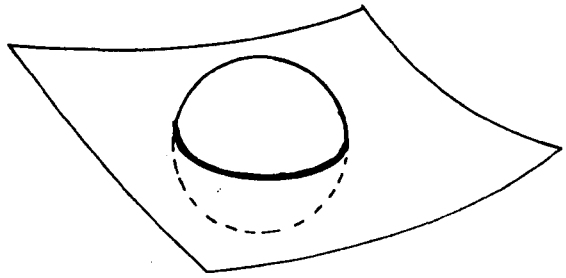
fig. 9.16)* A distorted rectangular grid.

space. The lines can then be coloured separately according to the Gaussian curvature. In the *CATIA* system elliptic points (with positive Gaussian curvature) were coloured blue and hyperbolic points were coloured green.

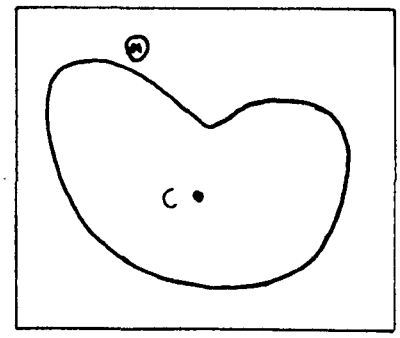
Rectangular grids are fine for drawing simple surfaces but many disadvantages are found when they are used to draw focal surfaces. These are mainly due to the fact that the grid becomes very distorted (fig 9.16)*). Certain parts become elongated while other parts are closely bunched together, making it difficult to see what is happening. The attention is drawn to the elongated parts which are merely features of the parametrization and not actually relevant. Furthermore the presence of grid lines may be misleading. Important parts of the surface may be missed out as the grid may not reach far enough. Elongated parts can be removed by clipping with a sphere, i.e. only drawing those parts contained inside the sphere. This does not cure the other problems.

§9.11 Polar grids

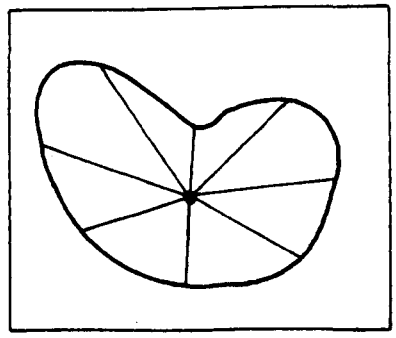
Many of the problems encountered in the previous section can be solved by using a grid centered on a point c define by polar coordinates. First we define a sphere centered on $f(c)$ in the image and find its intersection with the surface (fig 9.17a). This defines a curve Θ in parameter space (fig 9.17b) which loops round



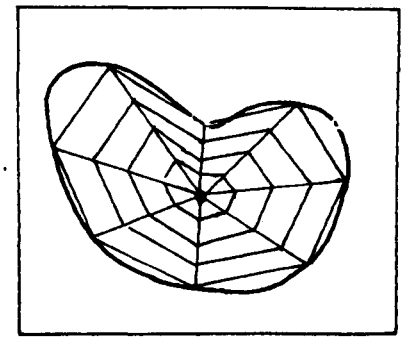
a) intersection of a sphere and a parametrized surface



b) pre-image, Θ, of line of intersection in parameter space



c) radial lines



d) final grid

fig. 9.17) Calculating a polar grid

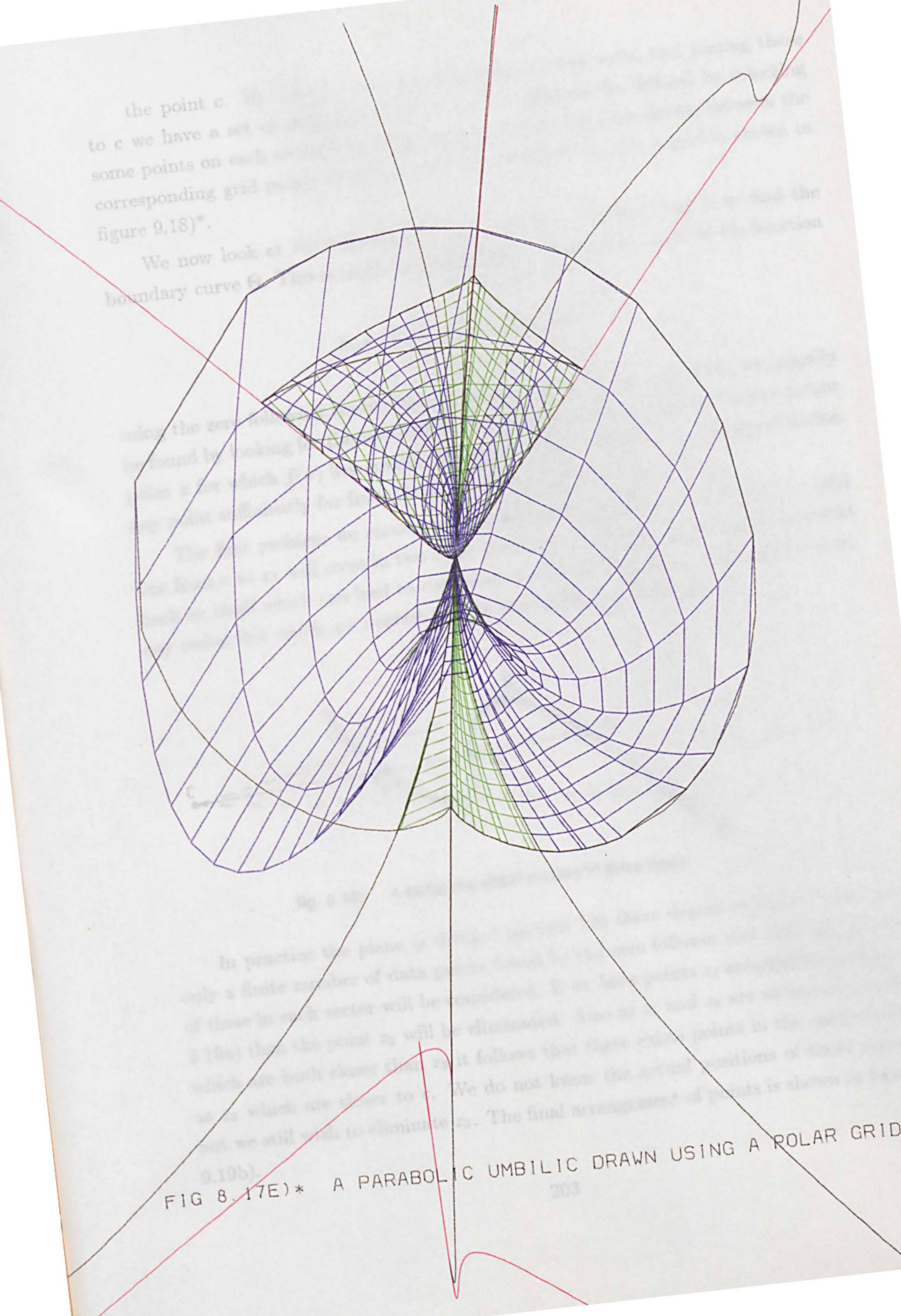


FIG 8.17E)* A PARABOLIC UMBILIC DRAWN USING A POLAR GRID

the point c. By taking a set of points lying on this curve and joining these to c we have a set of radial lines (fig 9.17c). A grid can be defined by selecting some points on each of these lines (fig 9.17d). Finally lines are drawn between the corresponding grid points on the surface. An example of such a grid is shown in figure 9.18)*.

We now look at the algorithm in more detail. The first step is to find the boundary curve Θ. This is easily accomplished by finding the zeros of the function

$$g_{c,r}(x) = ||f(c) - f(x)||^2 - r^2,$$

using the zero following routine discussed previously. A starting point can usually be found by looking for the zeros of this function along a line from c to some distant point x for which $f(x)$ lies outside the sphere. For a Monge form parametrization any point sufficiently far from the origin will satisfy this condition.

The first problem we encounter is illustrated by figure 9.18. Here the radial line from c to x_1 will cross in two other points x_2, x_3. The grid will appear to fold back on itself which can lead to confusing results. If we only take the points along any radial line which are nearest to c we can overcome this problem.

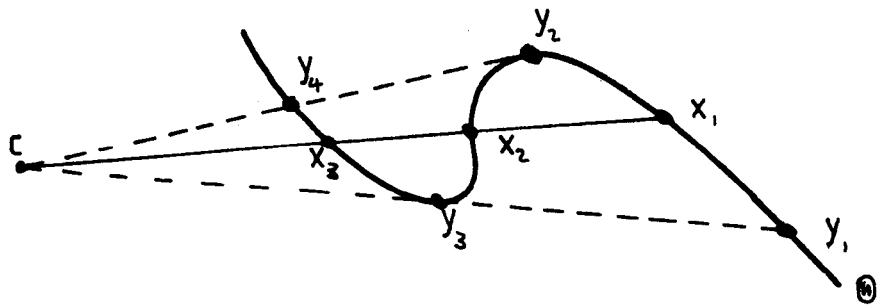


fig. 9.18) A radial line which crosses Θ three times

In practice the plane is divided up into 120 three degree sectors. There are only a finite number of data points found by the zero follower and only the closest of these in each sector will be considered. If we have points z_i arranged as in figure 9.19a) then the point z_2 will be eliminated. Also as z_4 and z_5 are successive points which are both closer than z_3 it follows that there exists points in the same sector as z_3 which are closer to c. We do not know the actual positions of these points but we still wish to eliminate z_3. The final arrangement of points is shown in figure 9.19b).

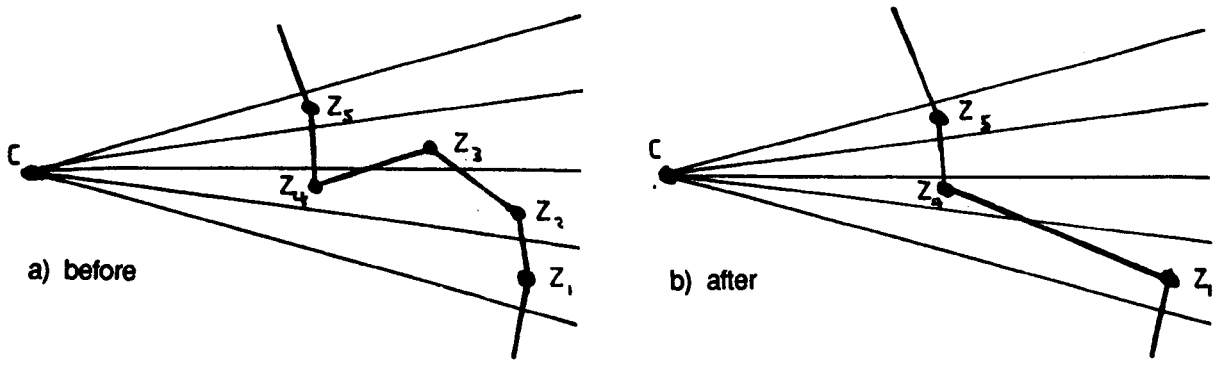


fig. 9.19) Eliminating points from sectors

This algorithm is fairly crude and may not achieve the optimum selection of points. However it has the advantages that it is simple and can be implemented with only one pass through the list. As long as the overlapping parts are removed it is not too important that some information is lost about the boundary.

The next step is to reduce the number of points round the boundary so that the grid is not too dense. This can be achieved by measuring the distance round the boundary in the image and then choosing points which are approximately a fraction of this distance apart. Whenever there is an overlapping situation like, figure 9.18), we will include some of the points with angles close to those of the turning points (y_2, y_3 in fig. 9.18). Here a turning point means a point where the value of the angle has a local extremum when moving along the curve Θ. This helps counteract the loss of information that occurred in the previous step. We now have m points a_i on the curve Θ and an $m \times n$ grid of points $\{x_{ij}\}$ can be constructed with

$$x_{ij} = a_i + \frac{j-1}{n}(x - a_i).$$

The grid does not extend right to the centre c, but the surface is often chosen so that it is highly singular here (for example an umbilic centre) and if drawn it might be misleading. The grid points can then be joined up to produce a circular grid.

One major problem is that the grid points do not actually lie on the ridges or sub-parabolic lines so when drawn on the focal surface the grid lines will miss the ribs or parabolic lines. This is particularly bad for ribs where the focal surface folds back on itself. In fig 9.20a)* the red lines represent the rib and the surface as draw does not actually touch it. If we find the intersection of the ridge with the straight line joining two grid points A, B we will find a point C (fig 9.21). Plotting this

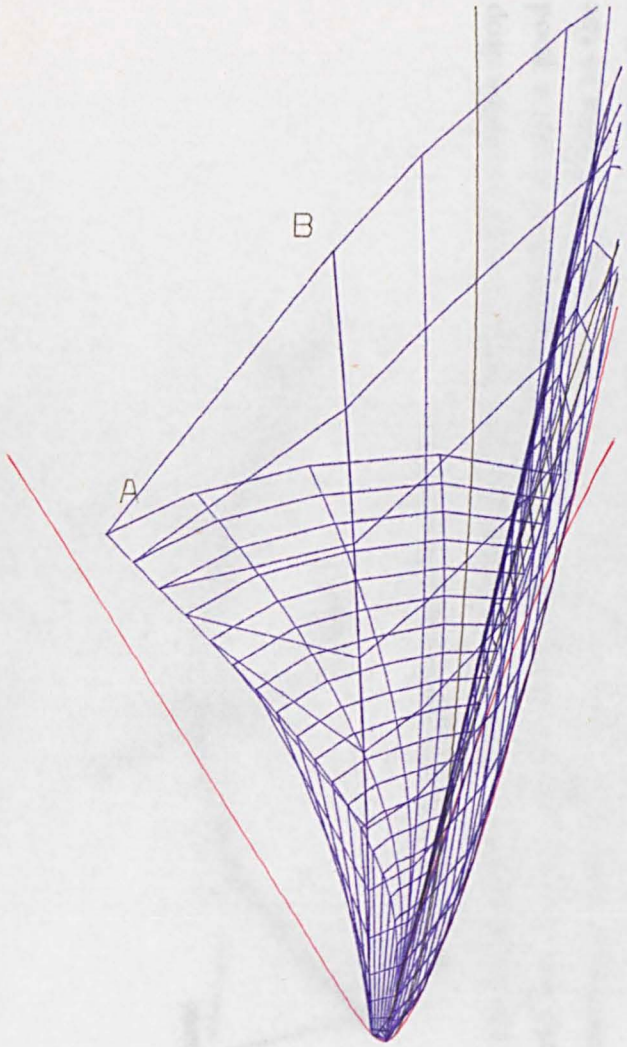


FIG 9.20 A) A GRID WITH POINTS ON RIB MISSING

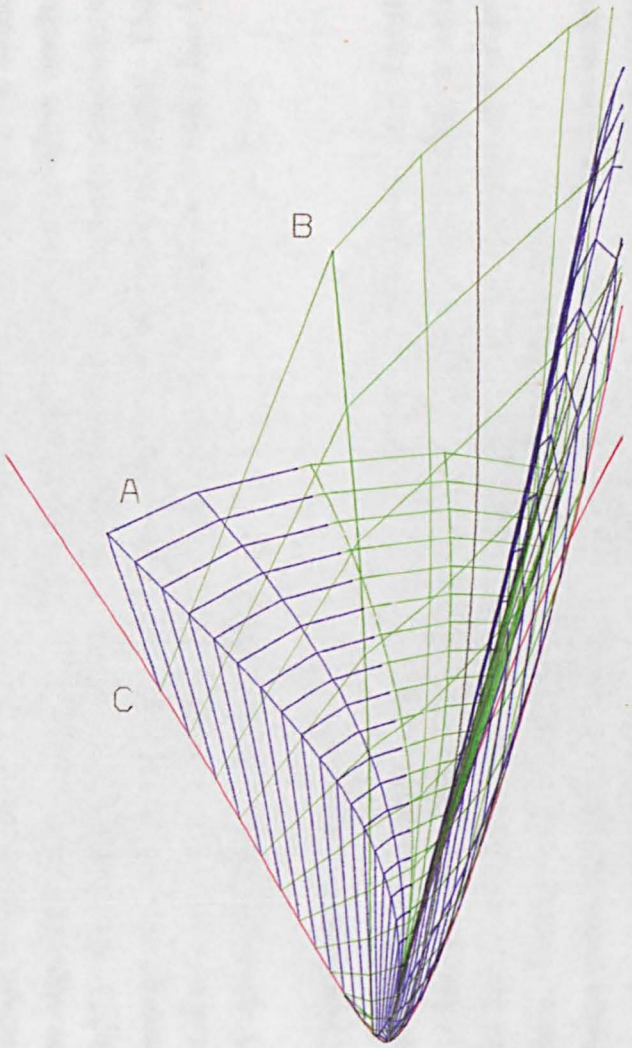


FIG 9.20 B) POINT C FOUND BY INTERSECTION OF RIDGE AND STRAIGHT LINE

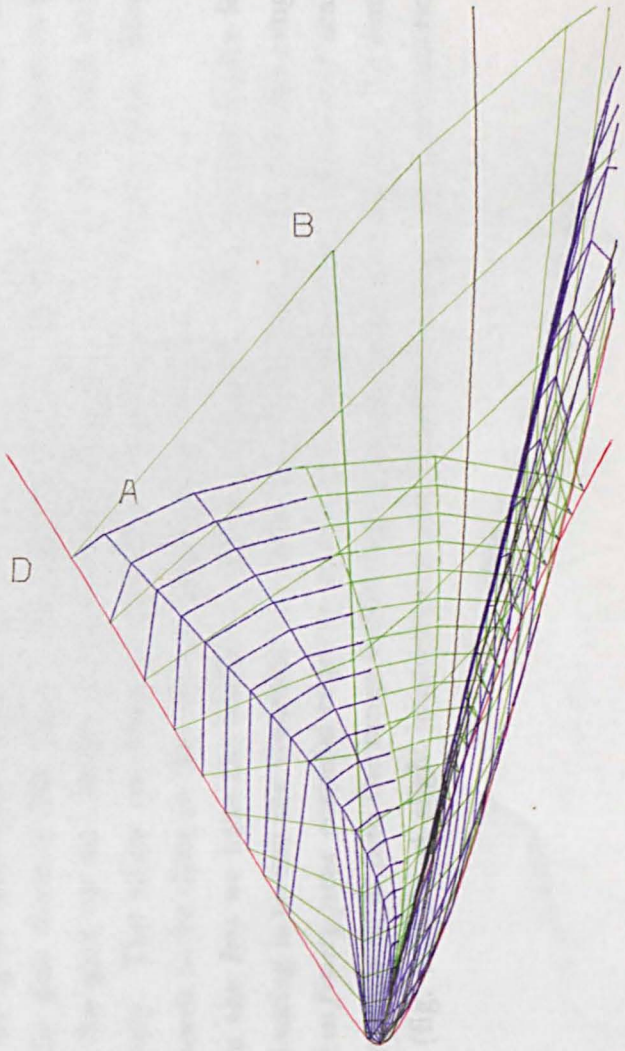


FIG 9.20 C) POINT D FOUND BY INTERSECTION WITH A SPHERE

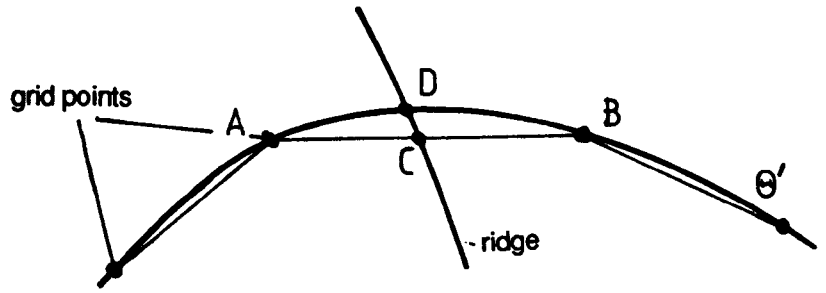


fig. 9.21) Intersection of ridge with grid lines

point on the focal surface will give a strange result (fig 9.20b)*) with a sudden drop as the grid line crosses the rib. This occurs because a straight line is not a good approximation to theoretical shape of the grid line. A better approximation to the shape is to assume that the curve is actually the pre-image Θ' of the intersection of the surface with a sphere. The intersection of Θ' with a ridge can then be found using the algorithm discussed in §9.7. Using this point D gives much better results (fig 9.20c)*). An alternative which has yet to be implemented is to find a polynomial curve through the grid points and find the intersection of this with the ridge. This problem does not occur for rectangular grids or radial lines, where a straight line is an exact representation of the grid lines.

§9.12 Creating *Nets* and *Tmeshes*

Both of the previous types of grid can be adapted to create a *net* or a *tmesh*. One problem encountered is deciding what should happen when a ridge or sub-parabolic line is crossed. The simplest solution is just to ignore the ridge and create the surface. This is what has so far been implemented on the *Iris*.

If extra points lying on the ridge are added to the grid points we will no longer have the correct arrangement to use the software. For instance in a *net* each region inside the grid lines must have four sides. Adding a segment of the ridge will normally give three or five (fig. 9.22). One way of simplifying the problem is to divide the grid up into strips which are one unit wide and work with each strip separately. The strips can then be cut along the ridge lines to form *nets*. More work needs to be done in this area to improve the solution.

On the *Iris* we have the ability to colour the surface. The obvious choice for the colouring is to use the Gaussian curvature K. However the values of this range from $\pm infinity$ near cuspidal edges to $\pm zero$ near parabolic lines. If we just scale the curvature to fit into the range 0–255 then most of the points will have the same colour (fig. 9.23)*). A slight improvement can be made by applying some function

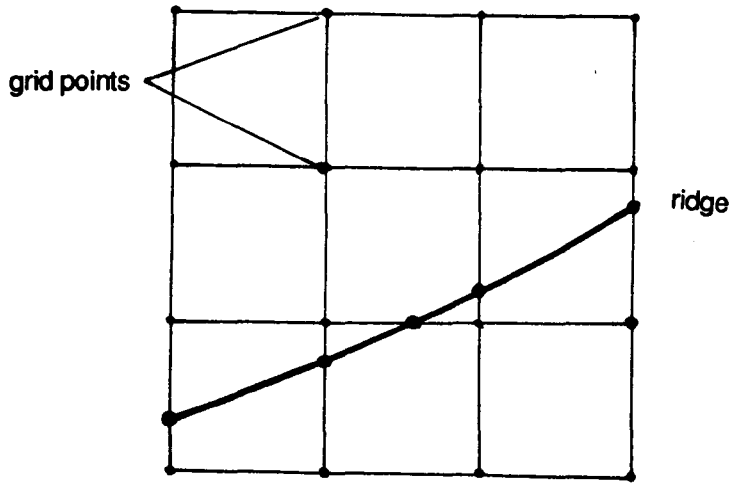
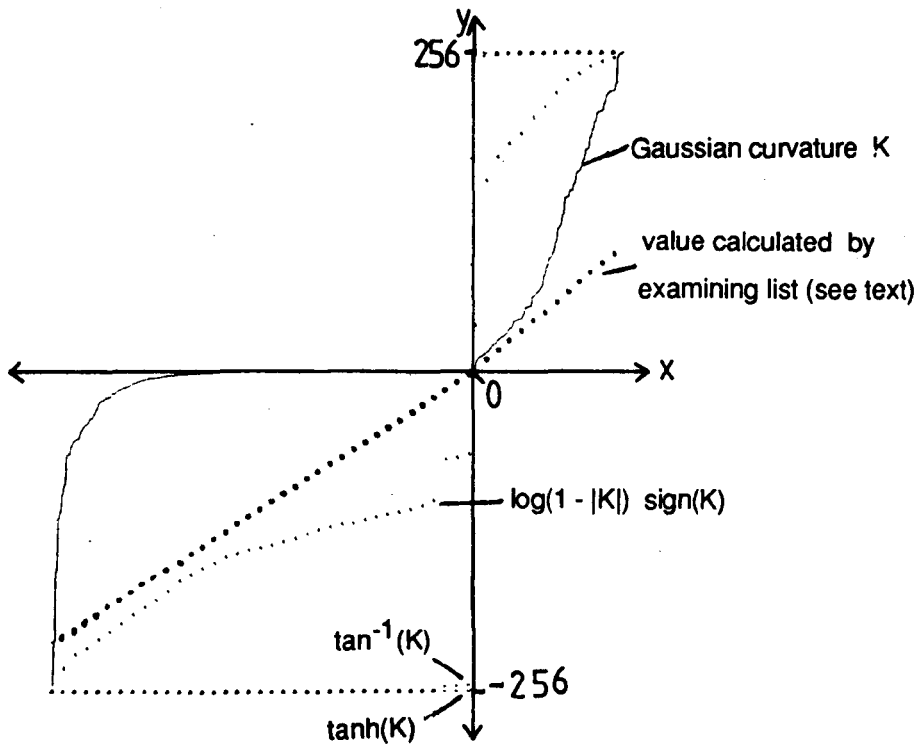


fig. 9.22) Adding points on a ridge to a set of grid points



+ve x-axis the position of each point in a list ordered by Gaussian curvature, positive values of Gaussian curvature

-ve x-axis the position of each point in a list ordered by Gaussian curvature, negative values of Gaussian curvature

+ve y-axis the value of each function scaled to fit range 0--256, positive values of Gaussian curvature

-ve y-axis the values of each function scalled to fit range -256-0, negative values of Gaussian curvature

fig 9.23)* The values of various functions of the Gaussian curvature

to the curvature and using the resultant values. The following functions were tried: $\tan^{-1}(K)$, $\tanh(K)$ and $\text{sign}(K)\ \log(|K|+1)$.

Much better results were obtained by sorting through the data beforehand. First the values of the curvature were sorted into a two lists with increasing values of curvature, one for negative values and one for positive values. Each list is then divided up into 256 divisions with approximately the same number of data points in each division. Finally we change the original unsorted data: the value of the curvature is replaced by the number of the division that value lies in. This number is then used for the colour value. Very pleasing results were obtained from this method as can be seen from the various figures in chapter 8. The use of a Sigmoid function could perform the same function.

The possibility of drawing shaded models has yet to be explored. To draw shaded pictures the normals to the surface must be calculated. For surfaces these are just the principal directions. As explained in §9.4 we can not find a continuous field of unit normals around an umbilic. So problems are anticipated.

§9.14 Other methods

A few other methods of representing the surfaces were tried. One is to draw the intersection of the focal surface with spheres of different radii (fig 9.24)*). This is computationally very expensive and the initial results were not very satisfactory.

An other method is to draw the **raised lines of curvature**. A line of curvature is a curve defined by the integral equation $r = \alpha p$ where r is the tangent to the curve, p is a principal direction and α is some real number (see §8.3). A raised line of curvature is image of this curve on one of the focal sheets. From §7.3 the tangent to a raised curve is

$$\underline{T} = df<p> = \beta \underline{n}$$

where $\underline{n}$ is normal to the original surface and $\beta = \alpha\ d^3V<p^3>/\kappa_p$ is a real number. Differentiating this along the curve gives

$$d\underline{T}<p> = \beta d\underline{n}<p> + d\beta<p>\underline{n}.$$

So as $d\underline{n}<p>\cdot\underline{Q} = -d^2s<pq> = 0$ the geodesic curvature $dT<p>\cdot\underline{Q}$ of this curve is zero. Hence when raised onto the focal sheet corresponding to its principal direction, we have a geodesic. It is also possible to raise the curve onto the other sheet these will not normally give geodesics but do help to illustrate the surface (fig. 9.25)*). As they form geodesics the raised lines of curvature are an interesting area of study in their own right.

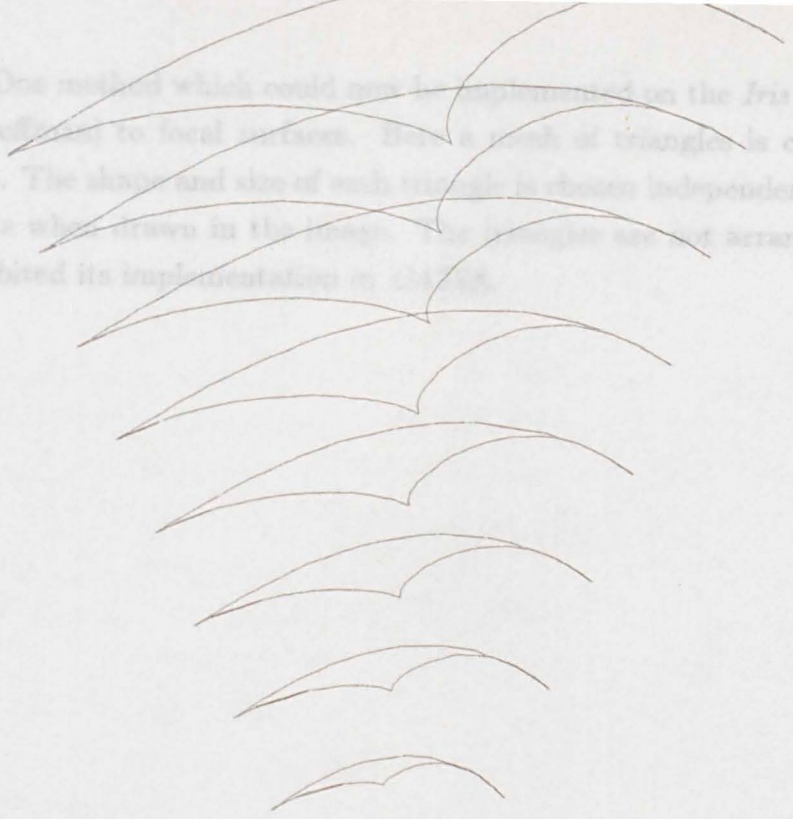


FIG 9.24) THE INTERSECTION OF A FOCAL SURFACE WITH SPHERES OF DIFFERENT RADI

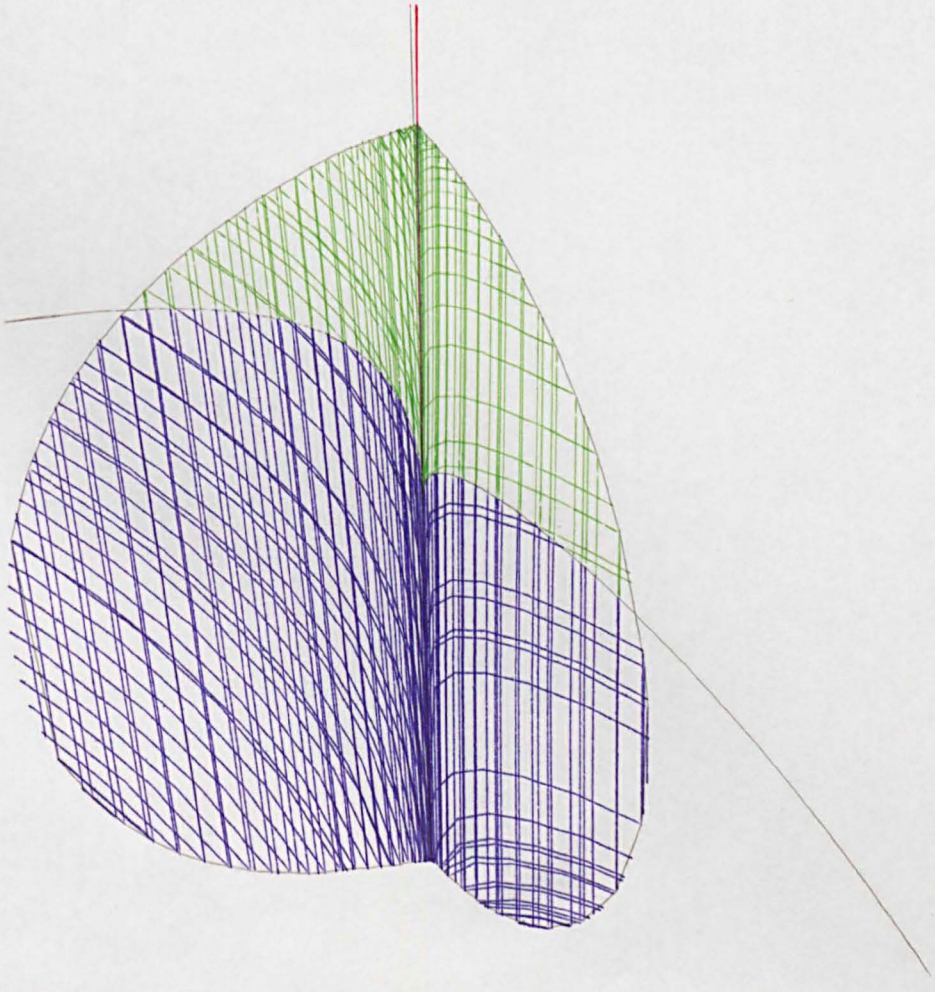


FIG 9.25) THE RAISED LINES OF CURVATURE ON A FOCAL SURFACE

One method which could now be implemented on the *Iris* is to adapt the work of [Hoffman] to focal surfaces. Here a mesh of triangles is created in parameter space. The shape and size of each triangle is chosen independently to give optimum results when drawn in the image. The triangles are not arranged regularly which prohibited its implementation in *CATIA*.

References

[**Arnold**] V.I.Arnold, Catastrophe theory, *Springer, Berlin,* (1984).

[**Blum**] H.Blum, Biological Shape and Visual Science (Part I), *J. Theor. Biol.* **38**, (1973), 205–287.

[**Brady–Asada**] M.Brady, H.Asada, Smooth Local Symmetries and their Implimentation, *Int. J. of Robotics Research,***Vol 3, No 3**, 36–61.

[**Brady**] M.Brady, Criteria for representation of shape, *in Human and Machine Vision, J.Beck, B.Hope, A.Rosenfeld Eds, Academic Press,* (1983), 39–84.

[**Bruce–Fidal**] J.W.Bruce, D.L.Fidal, On binary differential equations and umbilics, *Proc. Roy. Soc. Edinburgh,* **111A**, (1989), 147–168.

[**Bruce–Giblin–Gibson**] J.W.Bruce, P.J.Giblin, C.G.Gibson, Symmetry Sets, *Proc. Roy. Soc. Edinburgh,* **104A**, (1985), 163–186.

[**Bruce–Giblin–1**] J.W.Bruce, P.J.Giblin, Growth, motion on 1-parameter families of Symmetry Sets, *Proc. Roy. Soc. Edinburgh,***I04A**, (1986), 179–204.

[**Bruce–Giblin–2**] J.W.Bruce, P.J.Giblin, Curves and singularities, *Cambridge University Press,* (1984).

[**Bruce–Wilkinson**] J.W.Bruce, T.C.Wilkinson, Folding Maps and Focal Sets, *To appear Proc. Warwick Singularities Symp.,* (1989).

[**Canny**] J.F.Canny, Simplified Voronoi diagrams, *Proc. ACM symposium on Compuiational Geometry,* (1987).

[**Catia**] Catia version 3.0 reference manuals, *Dassault Systems, France.*

[**Darboux**] G.Darboux. Leçons sur la thèorie générale des surfaces, *Paris: Gauthiers Villars,* **Vol 4**, (1896).

[**do Carmo**] M.P.do Carmo, Differential geometry of curves and surfaces, *Pentice-Hall Inc, Englewood Cliffs, N.J.,* (1976).

[**Dobkin–Thurston**] Dobkins, Thurston, The Geometry of Circles: Voronoi diagrams, Mobius transformations, Convex Hulls, Fortune's algorithim, the Cut Locus and parameterization of shape, *Notes for lecture course, Princeton Univ,* (1986).

[**Giblin–Banchoff**] P.J.Giblin, T.F.Banchoff, On the geometry of piecewise circular curves, *Preprint,* (1990).

[**Giblin–Brassett**] P.J.Giblin, S.A.Brassett, Local Symmetry of Plane Curves, *Ammerican Math. Monthly,***Vol 92, No 10,** (1985), 689–707.

[**Giblin–Tari**] P.J.Gibin, F.Tari, Local Reflectional and Rotational Symmetry in the Plane, *To appear Proc. Warwick Singularities Symp.,* (1989).

[**Hannay**] M.V.Berry, J.H.Hannay, Umbilic points on Gaussian random surfaces, *J. Phys. A,* **10**, (1977), 1809–1821.

[**Hoffman**] M.J.Callahan, D.Hoffman, J.T.Hoffman, Computer Graphics Tools for the Study of Minimal Surfaces, *Communications of the ACM,***Vol 31, No 6,** (1988).

[**Iris**] Iris 4D GT Graphics Library Reference Manual, Silicon Graphics.

[**Koenderink**] J.J.Koenderink, Solid Shape, *MIT press,* (1990).

[**Leyton**] M.Leyton, Symmetry-Curvature Duality, *Computer Vision, Graphics, and Image Processing,***38** (1987), 327–341.

[**Markatis**] S.Markatis, Some generic phenomena in famlies of surfaces in $\mathbf{R}^3$, *Ph.D. thesis, University of Liverpool,* (1980).

[**O'Neill**] B.O'Neill, Elementary differential geometry, *Academic Press, New York,* (1966).

[**Porteous–1**] I.R.Porteous, Normal singularities of surfaces in $\mathbf{R}^3$, *Proc. Sympos. Pure Math.* **40**, (1983), 379–393.

[**Porteous–2**] I.R.Porteous, Probing Singularities, *Proc. Sympos. Pure Math.* **40** (1983), 395–406.

[**Porteous–3**] I.R.Porteous, *Notes for lecture course, University of Liverpool,* (1990).

[**Scott–Turner–Zisserman**] G.L.Scott, S.C.Turner, A.Zisserman, Using a mixed wave / diffusion process to elicit the Symmetry Set, *Image and Vision Computing,* **Vol 7, No 1**, (1989).

[**Sotomayer–Guttierrez**] J.Sotomayer, C.Guttierrez, Structually stable configurations of lines of principal curvature, *Astérisque,* **98-99**, (1982).

[**Tari**] F.Tari, On some applications of singularity theory to differential geometry and shape analysis, *Ph.D. thesis University of Liverpool,* (1990).

[**Wall**] C.T.C.Wall, Generic properties of generic differentiable manifolds, *Springer lecture notes in mathematics,***597**, (1977),707–774.