:chanism model for threshold spatial vision.

summation with changing locus in the visual

chapter 8

*Shape decompositions for visual recognition: the role of transversality**

B. M. Bennett
Department of Mathematics
University of California
Irvine, CA 92717

D. D. Hoffman
School of Social Sciences
University of California
Irvine, CA 92717

1. INTRODUCTION

Most theories of shape recognition agree that to recognize a complex shape it is useful to decompose the shape into simpler parts. The reasons are straightforward (Hoffman & Richards 1984). One never sees all of an opaque object at once; certainly its back is not visible, and even its front may be partially occluded by objects interposed between it and the viewer. So unless one can afford the luxury of seeing an object in its entirety before recognizing it (perhaps by walking around it), one must recognize objects from only partial information. In addition, some objects have moveable parts, such as arms or fingers, which allow them to assume many configurations. Decomposing such objects into appropriate parts, thereby decoupling configuration from other aspects of their shapes, can make easier their recognition. Finally, a

*We thank D. Mumford, W. Richards and R. Riley for useful comments. This material is based upon work supported by the National Science Foundation under Grant No. IST-8413560, and by a contract to D. Hoffman from the Office of Naval Research, Psychological Sciences Division, Engineering Psychology Group.

classification or description of parts, if the parts are appropriately chosen, is likely to be simpler than a classification of arbitrary shapes, and indeed should contribute to such a more general classification.

Although most theories agree that parts, in principle, are useful for recognition, they often disagree about how parts should be defined. This despite the widely acknowledged constraints that the parts on an object, however the parts are defined, should not in general change with minor changes in viewing geometry, i.e., with minor changes in the relative positions of the object and viewer, nor should the parts change with minor changes in overall size of the object.

There are two distinct approaches to the problem of part definition. The first, and by far most common, defines parts by their shapes; the second defines parts by their boundaries. The most typical parts proposed by partisans of the first approach, henceforth the *primitive-based* approach, are cylinders, cones, spheres, and polyhedra. Cylinders and cones are quite useful parts for representing the shapes of animals since many of their limbs are roughly cylindrical; polyhedra, on the other hand, do quite nicely for many buildings, books, and some furniture. Once the primitive part shapes are stipulated, it then remains to determine how to find these parts in complex objects, how to represent the more metrical aspects of the primitive parts (e.g., their length and width), and how to assign predicates of spatial relationships among parts (e.g., above, inside, to the right of).

To date there are but two theories representative of the second, or *boundary-based*, approach. Koenderink and van Doorn (1980, 1982) were the first to suggest that parts should be defined by their boundaries (though they were studying shading, not shape recognition); they propose that the appropriate boundaries are parabolic contours, i.e., contours on a surface where the Gaussian curvature is zero. Such contours possess several attractive properties. For instance, parabolic lines do not intersect and are always closed contours. By choosing them as part boundaries, Koenderink and van Doorn find that, on smooth surfaces of genus zero (no holes), there are only four qualitative classes of parts, which they call humps, dimples, furrows, and ridges.

Hoffman and Richards (1984) also suggest that parts should be defined by their boundaries rather than by a prespecified set of primitive shapes; they propose that part boundaries, instead of being defined by parabolic curves, should be defined by contours of negative minima of principal curvatures along lines of curvature, or in some cases, by contours of positive maxima of principal curvatures along lines of curvature. Hoffman and Richards' proposal will be discussed and extended further in this paper.

What are the relative merits of the boundary-based and primitive-based approaches to the problem of part definition? If one wants only to recognize a limited class of objects, say animals or aircraft, then the primitive-based approach is quite satisfactory. If, on the other hand, one wants a general pur-

pose shape recognizer, then the primitiv
simple reason that most shapes—human
merely of cylinders, cones, spheres, pol
And adding new primitives as needed to
hardly a way to build a principled th
approach, as exemplified by the parabo
Doorn, does give a part definition whicl
are guaranteed to exist on any smooth c
gle, and easily handled, exception) a
behaved partition of the surface. Once
specified, it becomes a differential geor
kinds of parts that can indeed arise. T
already done for compact smooth surfac
son the investigation need be restricted
faces. In this way a completely gen
definition and part description can be ob
approach confuses the problem of part c
part description, taking the latter to be tl

Although the boundary-based appr
Doorn has much to recommend it, the
rules and one must be careful, if one's
partitioning rule not primarily for n
relevance to the recognition task. (Koe
analysis of shading, not shape recogni
shortly the role of transversality in cons
Here we simply note that the human vi
parabolic lines in its definition of part
perception of parts which follow straigh
titioning rule are easily disconfirmed
Figures 1 and 2). According to the par
should not change on a surface when fi
abolic lines remains unchanged and (2)
surfaces since these surfaces have zero
ure 1, a cosine surface, disconfirms the
circular contours in the figure initially
successive ring-like parts. Now turn th
dotted circular contours no longer lie be
lie on top of them. In effect, the loca
change in figure and ground induced by
a cylindrical surface, which has zero G
a surface should have no parts (or infini
hypothesis because every point of the s
However we do see a small number of

ırts are appropriately chosen, is
itrary shapes, and indeed should
ı.

principle, are useful for recog-
ɔuld be defined. This despite the
ırts on an object, however the
: with minor changes in viewing
ıtive positions of the object and
ɔr changes in overall size of the

problem of part definition. The
ts by their shapes; the second
typical parts proposed by parti-
primitive-based approach, are
nders and cones are quite useful
since many of their limbs are
hand, do quite nicely for many
the primitive part shapes are
to find these parts in complex
aspects of the primitive parts
gn predicates of spatial relation-
ight of).
ɔresentative of the second, or
ın Doorn (1980, 1982) were the
y their boundaries (though they
; they propose that the appropri-
ontours on a surface where the
ɔssess several attractive proper-
sect and are always closed con-
Koenderink and van Doorn find
les), there are only four qualita-
imples, furrows, and ridges.
that parts should be defined by
d set of primitive shapes; they
ıg defined by parabolic curves,
minima of principal curvatures
contours of positive maxima of
Hoffman and Richards' proposal
ɔaper.
dary-based and primitive-based
If one wants only to recognize
rcraft, then the primitive-based
hand, one wants a general pur-

pose shape recognizer, then the primitive-based approach is inadequate for the simple reason that most shapes—human faces for instance—are not composed merely of cylinders, cones, spheres, polyhedra, or some combination of these. And adding new primitives as needed to handle new objects one encounters is hardly a way to build a principled theory. However, the boundary-based approach, as exemplified by the parabolic lines rule of Koenderink and van Doorn, does give a part definition which is completely general, for such lines are guaranteed to exist on any smooth compact surface (ovoids being the single, and easily handled, exception) and to provide a complete and well-behaved partition of the surface. Once the boundaries of parts have been so specified, it becomes a differential geometrical investigation to determine the kinds of parts that can indeed arise. This Koenderink and van Doorn have already done for compact smooth surfaces of genus zero, and there is no reason the investigation need be restricted from more complicated classes of surfaces. In this way a completely general, and principled, theory of part definition and part description can be obtained. In a sense, the primitive-based approach confuses the problem of part definition with the separate problem of part description, taking the latter to be the former.

Although the boundary-based approach taken by Koenderink and van Doorn has much to recommend it, there are many possible boundary-based rules and one must be careful, if one's goal is shape recognition, to choose a partitioning rule not primarily for mathematical convenience but for its relevance to the recognition task. (Koenderink & van Doorn's goal was an analysis of shading, not shape recognition.) In this regard, we will discuss shortly the role of transversality in constructing a definition of part boundary. Here we simply note that the human visual system does not appear to employ parabolic lines in its definition of parts, because two predictions about our perception of parts which follow straightforwardly from a parabolic lines partitioning rule are easily disconfirmed (the two exceptions are illustrated in Figures 1 and 2). According to the parabolic contours rule (1) part locations should not change on a surface when figure and ground reverse since the parabolic lines remains unchanged and (2) no parts should be seen on cylindrical surfaces since these surfaces have zero Gaussian curvature everywhere. Figure 1, a cosine surface, disconfirms the first prediction. Notice that the dotted circular contours in the figure initially appear to lie in the troughs between successive ring-like parts. Now turn the figure upside-down and note that the dotted circular contours no longer lie between the ring-shaped parts but rather lie on top of them. In effect, the location of the parts has changed with a change in figure and ground induced by inverting the surface. Figure 2 depicts a cylindrical surface, which has zero Gaussian curvature at every point. Such a surface should have no parts (or infinitely many parts) by the parabolic lines hypothesis because every point of the surface should lie on a part boundary. However we do see a small number of parts in this shape, which disconfirms

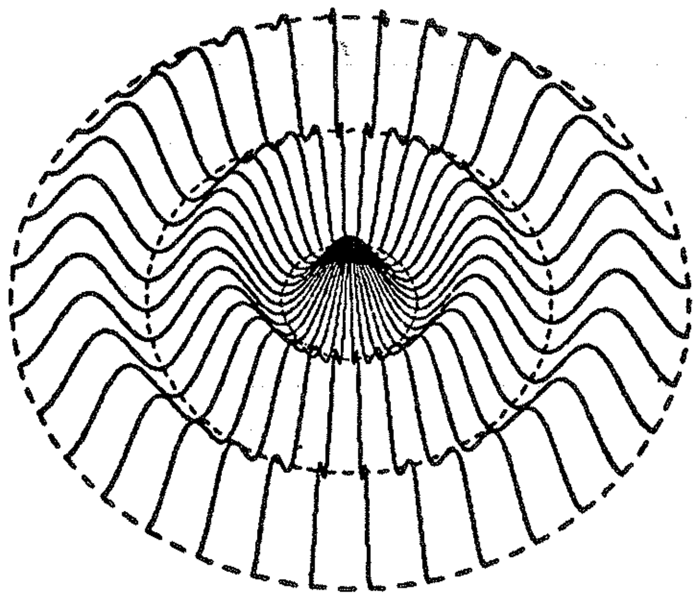


Figure 1. The cosine surface

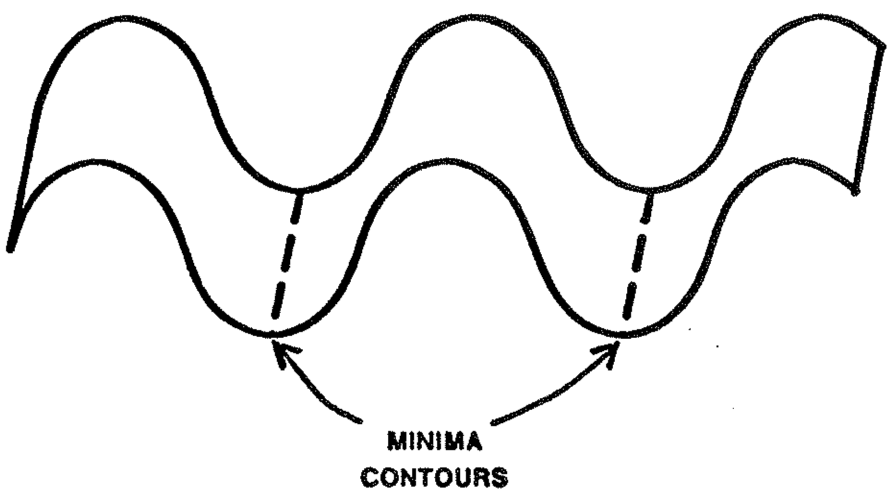


Figure 2. A cylindrical surface

the second prediction. N
hump-like parts with part
lines are drawn. Again, if
zation into parts can be se
within parts.

2. MOTIVATION OF

The motivation for the p
Hoffman and Richards (1
and arbitrarily shaped obj
Figure 3. Surely these tw
allow one of the objects t
thus forming a new comp
are good candidates for p
of points where the surfac
is a good candidate for the

Is there any special pr
be used to identify the lo
and thereby to identify the
two surfaces intersect the
means that the tangent pla
orientations at each point
that there is a discontinu
composite object at each
3). If the two original sur
when the discontinuity is
which is the case illustrate
original surface is remov
depression bounded by a
as is illustrated in Figure

This intuitive descripti
Consider two compact ob

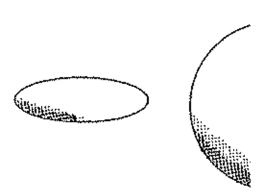
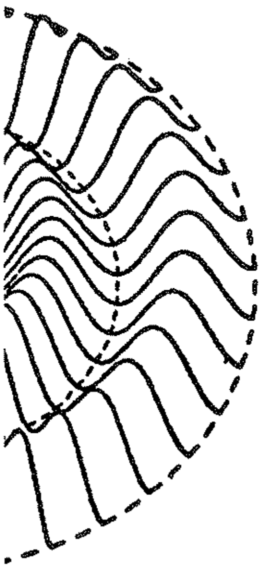
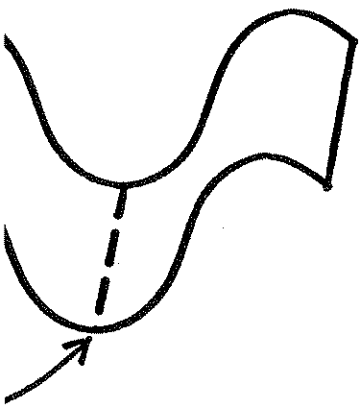


Figure 3. Transversal inte





the second prediction. Notice that this surface appears to be composed of hump-like parts with part boundaries located approximately where the dotted lines are drawn. Again, if the figure is turned upside-down a different organization into parts can be seen. The dotted lines no longer lie between parts but within parts.

2. MOTIVATION OF PARTITIONING RULES

The motivation for the partitioning rule proposed by Hoffman (1983) and Hoffman and Richards (1984) is roughly as follows. Consider two separate and arbitrarily shaped objects in a visual scene, as shown in the left half of Figure 3. Surely these two 3-D objects are separate parts of the scene. Now allow one of the objects to penetrate the other at some arbitrary orientation, thus forming a new composite object. Then certainly the two original objects are good candidates for parts of the resulting composite object, and the locus of points where the surface of the first object meets the surface of the second is a good candidate for the part boundary.

Is there any special property about the way two surfaces intersect that can be used to identify the locus of their intersection on the composite surface, and thereby to identify the boundary between the parts? Indeed there is: when two surfaces intersect they intersect transversally with probability one. This means that the tangent planes to the two intersecting surfaces are of different orientations at each point where the surfaces intersect. This implies further that there is a discontinuity of the tangent plane to the surface of the new composite object at each point along the contour of intersection (see Figure 3). If the two original surfaces are left together to form the composite surface when the discontinuity is *concave* at each point on the contour of intersection, which is the case illustrated in Figure 3. If, on the other hand, one of the two original surface is removed subsequent to penetrating the other, it leaves a depression bounded by a contour of *convex* discontinuity of the tangent plane, as is illustrated in Figure 4.

This intuitive description can be made more precise in the following way. Consider two compact objects, say obj_1 and obj_2, whose surfaces are given,

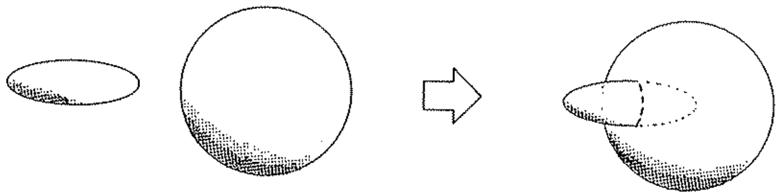


Figure 3. Transversal intersection leading to a protruding part

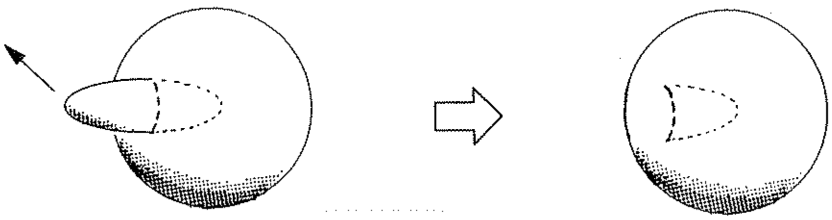


Figure 4. Transversal intersection leading to a intruding part

respectively, as the zero level sets of the two functions $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$. Here $\mathbf{x} = [x, y, z] \in R^3$. (A *level set* of a function, f, corresponding to some constant, c, is the set of all points, Q where $f(Q) = c$). The functions f_1 and f_2 are sometimes called "inside-outside" functions in the computer graphics literature (Barr 1983; Blinn 1982), because they can be used to define which points in R^3 are inside the corresponding object and which are outside:

$$\mathrm{obj}_1 = \{\mathbf{x} \in R^3 \mid f_1(\mathbf{x}) \leqslant 0\}$$

$$\mathrm{obj}_2 = \{\mathbf{x} \in R^3 \mid f_2(\mathbf{x}) \leqslant 0\}$$

That is, points in R^3 for which the inside-outside function is negative or zero constitute the object, whereas points for which the function is positive are outside. For instance, obj_1 might be a sphere defined by the function

$$f_1(\mathbf{x}) = x^2 + y^2 + z^2 - 1$$

Points for which f_1 is negative lie inside the sphere, points for which f_1 is zero constitute its surface, and points where f_1 is positive lie outside.

The new composite object formed by interpenetrating obj_1 with obj_2 and leaving the two together can be defined as the closed-set solid union of obj_1 and obj_2:

$$\mathrm{obj}_{\mathrm{new}} = \mathrm{obj}_1 \bigcup \mathrm{obj}_2 = \{\mathbf{x} \in R^3 \mid \mathbf{x} \in \mathrm{obj}_1 \text{ OR } \mathbf{x} \in \mathrm{obj}_2\}$$
$$= \{\mathbf{x} \in R^3 \mid f_1(\mathbf{x}) \leqslant 0 \text{ OR } f_2(\mathbf{x}) \leqslant 0\}.$$

The surface of the new composite object is then (Barr, 1983)

$$\mathrm{Surf}(\mathrm{obj}_1 \bigcup \mathrm{obj}_2) = \{\mathbf{x} \in R^3 \mid \mathbf{x} \in \mathrm{Surf}(\mathrm{obj}_1) \ \&\ \mathbf{x} \notin \mathrm{obj}_2 \text{ OR } \mathbf{x} \notin \mathrm{obj}_1 \ \&\ \mathbf{x} \in \mathrm{Surf}(\mathrm{obj}_2)\}$$

which can be expressed in te

Surf(obj$_1$ ⋃ obj$_2$

This new composite surfa
tinuity, consisting of points v
The new object formed b
ing obj$_2$ can be defined as th

obj$_{new}$ = obj$_1$ −

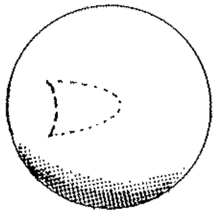
The surface of the resultir

Surf(obj$_1$ − c

which can be expressed in te

Surf(obj$_1$ − obj$_2$

This surface has, in gene
of points which satisfy f_1(x
Based on transversality,
discontinuity and some con
daries. Roughly, all contou
except those lying in the bo
discontinuity is a part bound
statement can be made p
differential geometry, but thi
One further step is need
surfaces such as the cosine
patch of surface having a co
is smoothed slightly (e.g., b
clear that the concave disco



ling part

ıctions $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$. Here ', corresponding to some con- = c). The functions f_1 and f_2 ıs in the computer graphics ' can be used to define which ınd which are outside:

$\leqslant 0\}$

$\leqslant 0\}$

utside function is negative or which the function is positive defined by the function

$- 1$

phere, points for which f_1 is s positive lie outside. enetrating obj_1 with obj_2 and closed-set solid union of obj_1

$\text{obj}_1 \text{ OR } \mathbf{x} \in \text{obj}_2\}$

$(\mathbf{x}) \leqslant 0 \text{ OR } f_2(\mathbf{x}) \leqslant 0\}.$

(Barr, 1983)

$\text{f}(\text{obj}_1) \ \& \ \mathbf{x} \notin \text{obj}_2$

$\text{j}_1 \ \& \ \mathbf{x} \in \text{Surf}(\text{obj}_2)\}$

which can be expressed in terms of inside-outside functions as

$$\text{Surf}(\text{obj}_1 \bigcup \text{obj}_2) = \{\mathbf{x} \in R^3 | f_1(\mathbf{x}) = 0 \ \& \ f_2(\mathbf{x}) > 0 \text{ OR } f_1(\mathbf{x}) > 0 \text{ OR } f_2(\mathbf{x}) = 0\}$$

This new composite surface has, in general, a contour of concave discontinuity, consisting of points which satisfy $f_1(\mathbf{x}) = f_2(\mathbf{x}) = 0$.

The new object formed by interpenetrating obj_1 with obj_2 and then removing obj_2 can be defined as the closed-set subtraction of obj_2 from obj_1.

$$\text{obj}_{\text{new}} = \text{obj}_1 - \text{obj}_2 = \{\mathbf{x} \in R^3 | \mathbf{x} \in \text{obj}_1 \ \& \ \mathbf{x} \notin \text{obj}_2\} = \{\mathbf{x} \in R^3; | f_1(\mathbf{x}) \leqslant 0 \ \& \ f_2(\mathbf{x}) > 0\}.$$

The surface of the resulting object is then (Barr, 1983)

$$\text{Surf}(\text{obj}_1 - \text{obj}_2) = \{\mathbf{x} \in \text{Surf}(\text{obj}_1) \ \& \ \mathbf{x} \notin \text{obj}_2 \text{ OR } \mathbf{x} \in \text{obj}_1 \ \& \ \mathbf{x} \in \text{Surf}(\text{obj}_2)\}$$

which can be expressed in terms of inside-outside functions as

$$\text{Surf}(\text{obj}_1 - \text{obj}_2) = \{\mathbf{x} \in R^3 | f_1(\mathbf{x}) = 0 \ \& \ f_2(\mathbf{x}) > 0 \text{ OR } f_1(\mathbf{x}) < 0 \ \& \ f_2(\mathbf{x}) = 0\}.$$

This surface has, in general, a contour of convex discontinuity, consisting of points which satisfy $f_1(\mathbf{x}) = f_2(\mathbf{x}) = 0$.

Based on transversality, then, we can take some contours of concave discontinuity and some contours of convex discontinuity to be part boundaries. Roughly, all contours of concave discontinuity are part boundaries except those lying in the bottom of a depression. And a contour of convex discontinuity is a part boundary only if it surrounds a depression. (This rough statement can be made precise and algorithmic using the language of differential geometry, but this is beyond the scope of this paper.)

One further step is needed to begin to define part boundaries on smooth surfaces such as the cosine surface of Figure 1. Consider what happens to a patch of surface having a concave discontinuity running through it if the patch is smoothed slightly (e.g., by draping a cotton sheet over it). Intuitively it is clear that the concave discontinuity will become a locus of very high curvature, in fact a locus of points which are negative extrema of surface curvature in a suitable sense (see Figure 5 and the partitioning rule stated below). To give this intuition a rigorous proof, however, requires dealing with some technicalities of a differential geometric nature, and so a complete and careful analysis must be made. In §3 of this chapter we develop the necessary analytic framework and prove a fundamental theorem of the behavior of curvature as smooth surfaces "approach" a transversal intersection (Theorem 6). The theorem says that near this intersection curve there are points of arbitrarily large negative curvature on these smooth surfaces. To complete a rigorous justification of the intuition it must be shown that these points form contours—as in the rule stated below—on members of the family of smooth surfaces, and that these contours approach the intersection curve. This will be done in a subsequent paper.

Negative Minima Partitioning Rule: Divide a surface into parts whose boundaries are contours consisting of points which are negative minima of principal curvature along a line of curvature.

The boundaries defined by this negative minima rule are derived in §4 for several classes of surfaces. The resulting boundaries are sketched in several figures in that section so that one can better appreciate the effect of the rule. To avoid confusion, it should be noted that although the minima rule employs lines of curvature in its definition of part boundary, *the part boundaries themselves are not, in general, lines of curvature.*

The negative minima rule provides boundaries for all parts except depressions. Depressions are delimited by positive maxima of the principal curvatures along their associated lines of curvature, which is the rule one obtains by smoothing convex discontinuities. A further set of rules, which will not be described here, determine when the negative minima rule is to be used and

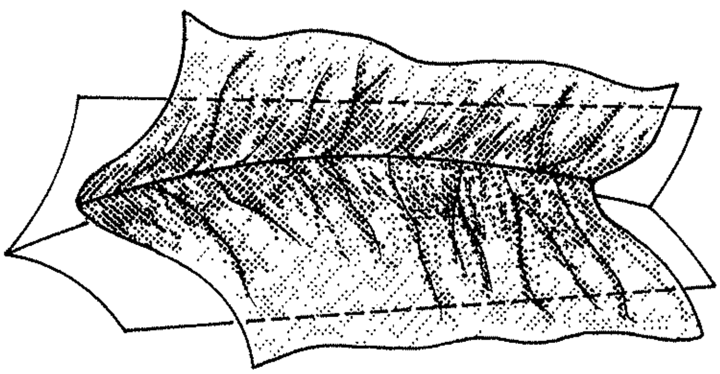


Figure 5. Smoothing a transversal intersection leading to negative minima of curvature, which in the figure occur along the curve in the shaded region

when, instead, the positive
negative minima of the princ
they correspond, roughly, to
imited by positive maxima
parts" since they correspond
This paper focuses entirel
of the principal curvatures.

3. DERIVATION OF PAR

In this section we will prove
faces leads to arbitrarily larg
case regardless of how the sn
poses a "smoothing" of a giv
converging to it in the C^2 sen

We begin with an examp
introduce in a concrete settin
the proof, we consider in det
sumes: (1) smoothing with a
mation.

3.1 Smoothing transversal i

One particularly simple exam
the two lines $y = 0$ and $x =$
6a. Consider the sets of p
$f(x, y) = xy = 0$, the so-
This level set is precisely the
on which either $x = 0$ or y

Representing this transvers
convenient representation f
$g(x, y) = xy - \epsilon$, where ϵ
of these functions $g(x, y)$ a
approach the case of transver
ϵ serves as a smoothing para
degree of smoothing. The pa
the family of level sets, wit
level set.

The curvature, $k(x, \epsilon)$, on
lae to be:

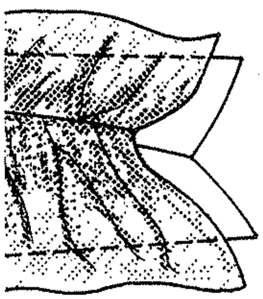
$k(x,$

e extrema of surface curvature
tioning rule stated below). To
:quires dealing with some tech-
nd so a complete and careful
we develop the necessary ana-
rem of the behavior of curva-
rsal intersection (Theorem 6).
urve there are points of arbi-
oth surfaces. To complete a
: shown that these points form
nbers of the family of smooth
ntersection curve. This will be

de a surface into parts whose
which are negative minima of

ima rule are derived in §4 for
daries are sketched in several
preciate the effect of the rule.
ough the minima rule employs
ary, *the part boundaries them-*

ies for all parts except depres-
aaxima of the principal curva-
which is the rule one obtains
set of rules, which will not be
ninima rule is to be used and



ling to negative minima of curva-
shaded region

when, instead, the positive maxima rule is to be used. Parts delimited by negative minima of the principal curvatures are called "positive parts" since they correspond, roughly, to various kinds of bumps on an object. Parts delimited by positive maxima of the principal curvatures are called "negative parts" since they correspond to depressions in an object.

This paper focuses entirely on positive parts delimited by negative minima of the principal curvatures.

3. DERIVATION OF PARTITIONING RULES

In this section we will prove that smoothing a transversal intersection of surfaces leads to arbitrarily large negative curvature. It turns out that this is the case regardless of how the smoothing is accomplished in C^2, i.e., for our purposes a "smoothing" of a given surface will be a sequence of smooth surfaces converging to it in the C^2 sense; for precise definitions see below.

We begin with an example intended to make plausible the claim and to introduce in a concrete setting several concepts used in the proof. Following the proof, we consider in detail two special cases of smoothing which it subsumes: (1) smoothing with a Gaussian and (2) smoothing by spline approximation.

3.1 Smoothing transversal intersections: An example

One particularly simple example of a transversal intersection is that formed by the two lines $y = 0$ and $x = 0$, i.e. by the x and y axes, as shown in Figure 6a. Consider the sets of points in the plane which satisfy the equation $f(x, y) = xy = 0$, the so-called "zero level set" of the function $f(x, y)$. This level set is precisely the desired two lines, because it is the set of points on which either $x = 0$ or $y = 0$.

Representing this transversal intersection by means of a level set leads to a convenient representation for smoothing. Consider the set of functions $g(x, y) = xy - \epsilon$, where $\epsilon \geqslant 0$. As ϵ approaches zero, the zero level sets of these functions $g(x, y)$ approach the zero level set of $f(x, y)$, i.e. they approach the case of transversal intersection, as shown in Figure 6b. In effect, ϵ serves as a smoothing parameter, with larger values of ϵ indicating a greater degree of smoothing. The parameter ϵ can also be thought of as an index into the family of level sets, with each value of ϵ uniquely associated with one level set.

The curvature, $k(x, \epsilon)$, on these level sets can be found by standard formulae to be:

$$k(x, \epsilon) = \frac{-2\epsilon x^{-3}}{(\sqrt{1 + \epsilon x^{-4}})^3}$$

For a particular choice of ϵ, i.e. for any particular member of the family of level sets, the curvature will have its greatest absolute value (and negative sign) at the point where the level set intersects the line $y = x$. This can be seen by noting the symmetry of the level sets about the line $y = x$ in Figure 6b. Now along this line we have that $x = y = \epsilon/x$, so that $x = \sqrt{\epsilon}$. Substituting this relation into the equation for curvature, and simplifying, we find that the negative minimum of curvature for the level set ϵ is

$$k_{\min}(\epsilon) = 1\sqrt{2/\epsilon}.$$

Now as $\epsilon \to 0$, i.e. as the level sets approach the singular level set, the minimum of curvature goes to $-\infty$. Thus we see intuitively that smoothing the transversal intersection by means of this family of level sets replaces the singular point with negative minima of curvature, as illustrated by Figure 6b.

3.2 Preliminaries on the curvature of level sets

The previous section demonstrated, for a simple example, that smoothing a transversal intersection leads to negative minima of curvature. In this section we begin the proof that this result holds for all transversal intersections and all smoothings.

We start by considering surface curvature. Curvature is a priori a property of a surface (or manifold) at a point. However, in most applications the surface in question is naturally defined as the level set of a function. Thus if f is

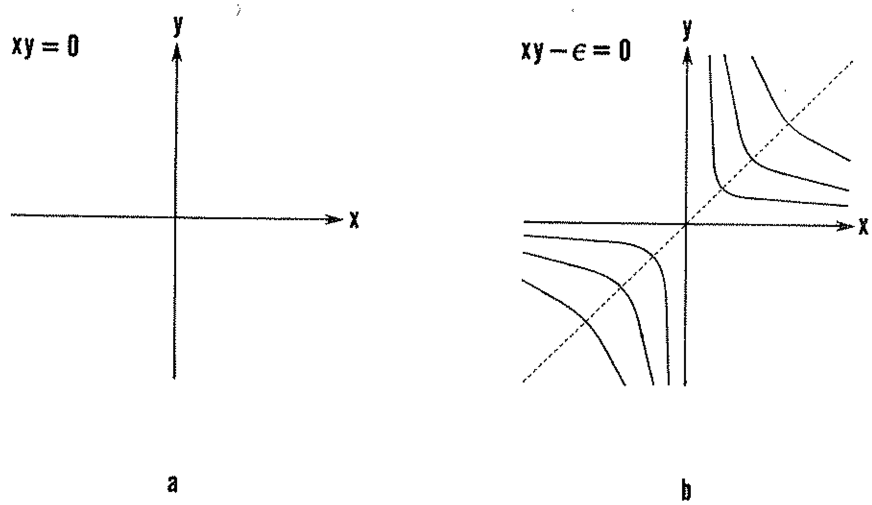


Figure 6ab. Smoothing of a transversal intersection by a parametrized family of level sets

a function on a domain D,
through P, i.e. the set

$M(f, P$

We note that there are
through P. For example
$f_1 = hf_2$, where h is a n
guity, our point of view h
defined by particular funct
practice, and also leads mo
which is our ultimate intere

The most general class o
able notion of curvature are
partial derivatives through
denotes the set of functions
to be the set of such functi
together with their derivat
extensions to the boundary
where we can define the me

$\|f_1 - f_2\|$

where ∂ ranges over all par
nate system) of order 0 thro

We now consider a dom
the level set $M(f, P)$. If ∇
smooth surface through P;
(x, y, z) so that $\nabla f(P)$ po
the origin, as shown in fig
C^2 function $g(x, y)$ on a n
near P, $M(f, P)$ is the g
$P = (0,0)$,

represents the so called "s
and its eigenvalues are by
denoted by $\kappa_1(f, P)$, $\kappa_2(f,$
we may assume that $g_{xy}(P$

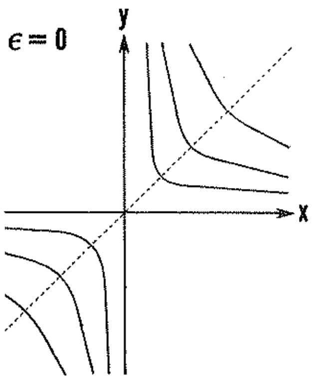
$g_{xx}(P)$
0 g

ar member of the family of
solute value (and negative
: line $y = x$. This can be
it the line $y = x$ in Figure
'x, so that $x = \sqrt{\epsilon}$. Substi-
:, and simplifying, we find
:l set ϵ is

the singular level set, the
: intuitively that smoothing
y of level sets replaces the
s illustrated by Figure 6b.

example, that smoothing a
f curvature. In this section
ansversal intersections and

ature is a priori a property
ı most applications the sur-
: of a function. Thus if f is



b

by a parametrized family of

a function on a domain D, and if $P \in D$, we can look at the level set of f through P, i.e. the set

$$M(f, P) = \{Q \in D \mid f(Q) = f(P)\}.$$

We note that there are many functions which have the same level set through P. For example the sets $f_1 = 0$ and $f_2 = 0$ are the same if $f_1 = hf_2$, where h is a nowhere vanishing function. In spite of this ambiguity, our point of view here will be that of surfaces in $\mathbf{R}^3$ as level sets defined by particular functions. This reflects the situations which arise in practice, and also leads most naturally to the study of variation of level sets, which is our ultimate interest in this section.

The most general class of functions on whose level sets there is a reasonable notion of curvature are the "C^2 functions", i.e. functions with continuous partial derivatives through the second order. If D is a domain in $\mathbf{R}^n$, $C^2(D)$ denotes the set of functions on D. If D is compact then $C^2(D)$ can be defined to be the set of such functions which are C^2 on the interior of D, and which, together with their derivatives through the second order, have continuous extensions to the boundary of D. With this definition $C^2(D)$ is a metric space, where we can define the metric, $\| \ \|_{C^2}$, as follows: Let $f_1, f_2 \in C^2(D)$. Then

$$\|f_1 - f_2\|_{C^2} = \sup_{P \in D}\{|\partial f_1(P) - \partial f_2(P)|\},$$

where ∂ ranges over all partial derivatives (with respect to some fixed coordinate system) of order 0 through 2.

We now consider a domain $D \subset \mathbf{R}^3$, and let $f \in C^2(D)$. For $P \in D$, we have the level set $M(f, P)$. If $\nabla f(P)$ (the gradient of f at P) $\neq 0$, $M(f, P)$ is a smooth surface through P; we can choose an orthogonal coordinate system (x, y, z) so that $\nabla f(P)$ points in the direction of the positive z-axis, and P is the origin, as shown in figure 7. By the implicit function theorem there is a C^2 function $g(x, y)$ on a neighborhood of the origin in the x, y-plane so that near P, $M(f, P)$ is the graph $z = g(x, y)$. The Hessian matrix of g at $P = (0,0)$,

$$\begin{bmatrix} g_{xx}(P) & g_{xy}(P) \\ g_{xy}(P) & g_{yy}(P) \end{bmatrix},$$

represents the so called "second fundamental form" of the surface M_f at P, and its eigenvalues are by definition the principal curvatures of $M(f, P)$ at P, denoted by $\kappa_1(f, P)$, $\kappa_2(f, P)$. Thus, after a suitable rotation of the xy-plane, we may assume that $g_{xy}(P) = 0$, so that the Hessian of g at P is now

$$\begin{bmatrix} g_{xx}(P) & 0 \\ 0 & g_{yy}(P) \end{bmatrix} = \begin{bmatrix} \kappa_1(f, P) & 0 \\ 0 & \kappa_2(f, P) \end{bmatrix}.$$

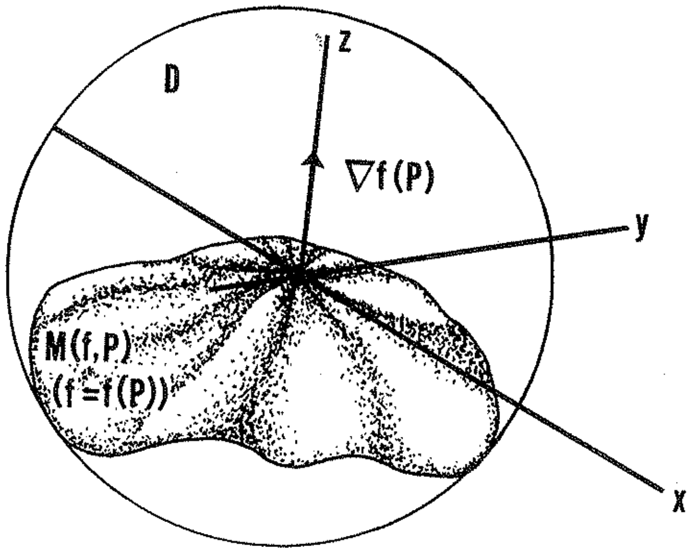


Figure 7. A level set M, with an orthogonal coordinate system centered at some point P of M

Now, from the relation $f(x, y, g(x, y)) =$ constant, we deduce

$$f_x(x, y, g) + f_z(x, y, g)g_x = 0, \quad f_x(x, y, g) + f_z(x, y, g)g_y = 0. \tag{3.1}$$

Using the fact that $\nabla f(P) = (0, 0, |\nabla f(P)|)$, i.e. that $f_x(P) = f_y(P) = 0$ and $f_z(P) \neq 0$, the equations (3.1) imply that $g_x(0, 0) = g_y(0, 0) = 0$. This, together with an additional differentiation of the equations (3.1) with respect to x and y yields:

$$f_{xx}(P) + f_z(P)g_{xx}(0, 0) = 0, \quad f_{yy}(P) + f_z(P)g_{yy}(0, 0) = 0,$$

that is,

$$\kappa_1(f, P) = f_{xx}(P)/f_z(P), \quad \kappa_2(f, P) = f_{yy}(P)/f_z(P),$$

and finally:

$$\kappa_1(f, P) = \frac{f_{xx}(P)}{\nabla f(P)}, \quad \kappa_2(f, P) = \frac{f_{yy}(P)}{\nabla f(P)}. \tag{3.2}$$

This expression depends on the particular x-y-z coordinate system which is associated as above to f and P. We want to express this same relation in a

form which is coordinat
level surfaces of f with
sider the Hessian of f it
on $\mathbf{R}^3$, which can be de
system, say $(u^1, u^2, u^3$
derivatives $fu_iu_i(P)$). T
elements, is independent
by tr$H(f, P)$.

Since in terms o
tr$H(f, P) = f_{xx}(P) +$

$$\kappa_1(f, P) +$$

Here z itself has an i
normal vector $\nabla f(p)/|\nabla$
$N(f, P)$, we may write

$$f_{zz}($$

(This means that in the
by a matrix A and $N($
Thus we may rewrite (3

Proposition 1: If f i

$$\kappa_1(f, P) + \kappa_2(f,$$

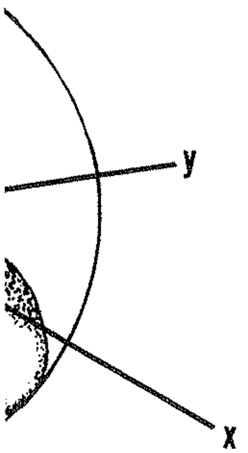
where $\kappa_i(f, P)$ are the
$H(f, P)$ is the Hessian
$M(f, P)$ at P, i.e. $N(f,$

Remark: The quan
"mean curvature" of the

3.3 Level sets with tra

As above, let D be a d
lowing property: the le
faces, B_1 and B_2, whi
shown in Figure 8.

The two smooth surf
locus" of M. The trans
at each point P_0 of S,
erates R^3, i.e. the tang
intersect in the tangent



dinate system centered at some

tant, we deduce

$+ f_z(x, y, g)g_y = 0.$ (3.1)

..e. that $f_x(P) = f_y(P) = 0$
t $g_x(0, 0) = g_y(0, 0) = 0.$
of the equations (3.1) with

$+ f_z(P)g_{yy}(0, 0) = 0,$

$= f_{yy}(P)/f_z(P),$

$\frac{f_{yy}(P)}{\nabla f(P)}.$ (3.2)

-z coordinate system which is
press this same relation in a

form which is coordinate free, so that we can compare the curvatures of the level surfaces of f with those of "nearby" functions. For this purpose we consider the Hessian of f itself at P, denoted $H(f, P)$. This is a quadratic form on $\mathbf{R}^3$, which can be defined intrinsically. In any given orthogonal coordinate system, say (u^1, u^2, u^3), it is represented by the matrix of second partial derivatives $fu_iu_i(P))$. The trace of this matrix, i.e. the sum of its diagonal elements, is independent of the particular coordinate system; we will denote it by $\text{tr}H(f, P)$.

Since in terms of the x-y-z system discussed above we have $\text{tr}H(f, P) = f_{xx}(P) + f_{yy}(P) + f_{zz}(P)$, the equations (3.2) imply:

$$\kappa_1(f, P) + \kappa_2(f, P) = \frac{\text{tr}H(f, P) - f_{zz}(P)}{|\nabla f(P)|}. \tag{3.3}$$

Here z itself has an intrinsic meaning as an axis in the direction of the unit normal vector $\nabla f(p)/|\nabla f(P)|$ to $M(f, P)$ at P. If we denote this unit vector $N(f, P)$, we may write

$$f_{zz}(P) = N(f, P)^t H(f, P) N(f, P).$$

(This means that in the given coordinate system, if the Hessian is represented by a matrix A and $N(f, P)$ by a column vector B, then $f_{zz}(P) = B^t AB$). Thus we may rewrite (3.3), and we get:

Proposition 1: If f is C^2 around $P \in \mathbf{R}^3$, and $\nabla f(P) \neq 0$, then

$$\kappa_1(f, P) + \kappa_2(f, P) = \frac{\text{tr}H(f, P) - N(f, P)^t H(f, P) N(f, P)}{|\nabla f(P)|},$$

where $\kappa_i(f, P)$ are the principal curvatures of the level set $M(f, P)$ at P, $H(f, P)$ is the Hessian form of f at P, and $N(f, P)$ is the unit normal to $M(f, P)$ at P, i.e. $N(f\ P) = \nabla f(P)/|\nabla f(P)|$.

Remark: The quantity $\kappa_1(f, P) + \kappa_2(f, P)$ is sometimes called the "mean curvature" of the surface $M(f, P)$ at P; we will denote it $\mu(f, P)$.

3.3 Level sets with transversal intersections

As above, let D be a domain in $\mathbf{R}^3$, and suppose that $\phi \in C^2(D)$ has the following property: the level set M: $\phi = 0$ consists locally of two smooth surfaces, B_1 and B_2, which intersect transversally along a smooth curve S, as shown in Figure 8.

The two smooth surfaces are called the "branches" of M; S is the "singular locus" of M. The transversality of the intersection of the branches means that at each point P_0 of S, the union of the tangent spaces to the branches generates R^3, i.e. the tangent planes are not parallel, and the two tangent spaces intersect in the tangent space to S:

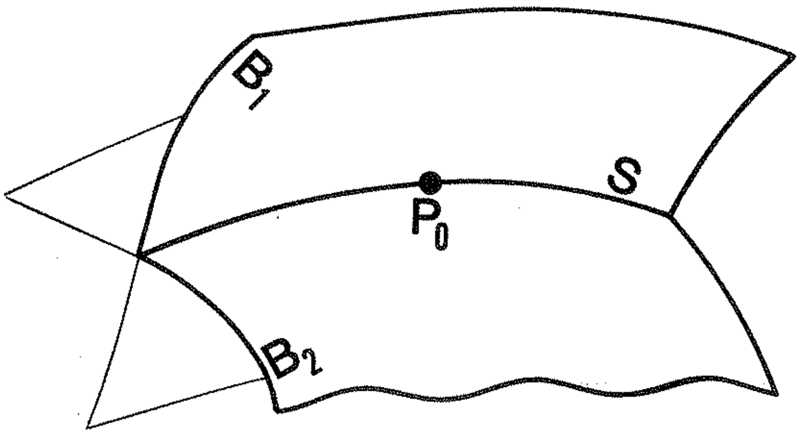


Figure 8. A level set consisting of two smooth branches intersecting transversally

$$T_{P_0}(B_1) \cap T_{P_0}(B_2) = T_{P_0}(S)$$

$$T_{P_0}(B_1) + T_{P_0}(B_2) = \mathbf{R}^3.$$

Given such a ϕ, and $P_0 \in S$, we will choose an orthogonal coordinate system (u, v, w) on $\mathbf{R}^3$ that $P_0 = (0, 0, 0)$, the u-axis is tangent to S, and the v and w axes are chosen as follows: Let $\mathbf{q}_1$ be a unit vector along the u-axis, i.e. $\mathbf{q}_1 \in T_{P_0}(S)$. For $\mathbf{i} = 1, 2$ let $\mathbf{r}_i \in T_{P_0}(B_i)$ with $\mathbf{r}_i \perp \mathbf{q}_1$ and $|\mathbf{r}_i| = 1$. Let $\mathbf{q}_3$ be $(\mathbf{r}_1 + \mathbf{r}_2)/|\mathbf{r}_1 + \mathbf{r}_2|$, i.e. $\mathbf{q}_3$ is a unit "bisector" of $\mathbf{r}_1$ and $\mathbf{r}_2$. Let $\mathbf{q}_2$ be $\mathbf{q}_3 \times \mathbf{q}_1$. Finally choose the coordinates v and w so that the positive v and w axes are in the $\mathbf{q}_2$ and $\mathbf{q}_3$ directions respectively. The picture is shown in Figure 9.

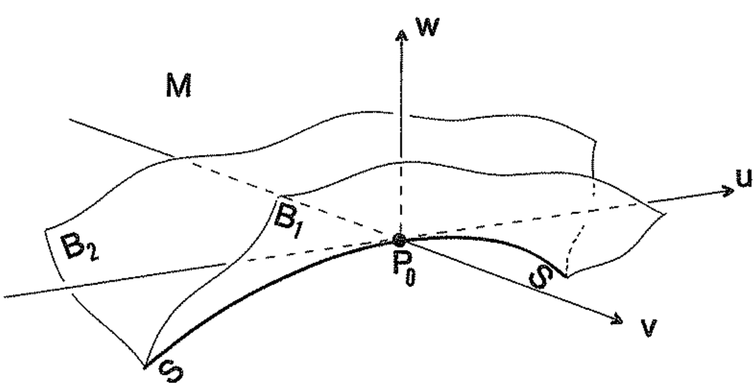


Figure 9. A canonical coordinate system associated with the level surface $M\colon \psi = 0$.

If we intersect this figure

Remark: Note that the ch
possible values of $\mathbf{q}_3$, any
opposite directions. Thus alte
or reverse their orientations.
of a given observer, however
which is "visible" to the obse

Let b_i be any C^2 function
v, w system $b_i(P_0) = b_i(0,$
have, near $P_0(0, 0, 0)$,

$$b_i(u, v, w) =$$

where $\alpha_i = \partial b_i/\partial u(P_0)$, β_i
a C^2 function whose order
values, and that of its first
tion locus S is the common
S at P_0 is the null space of
space of

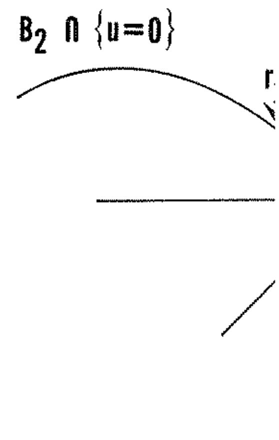
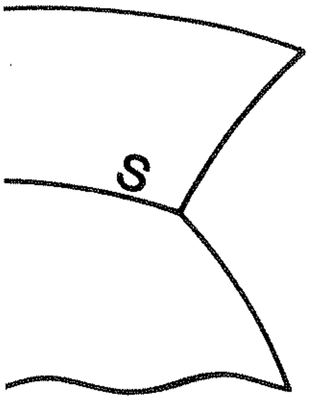


Figure 10. The intersection o

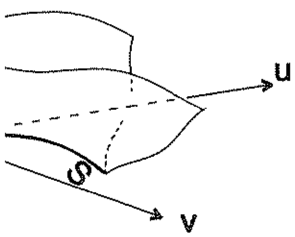


ınches intersecting transversally

$T_{P_0}(S)$

$= \mathbf{R}^3.$

e an orthogonal coordinate sys-
u-axis is tangent to S, and the v
unit vector along the u-axis, i.e.
ı $\mathbf{r}_i \perp \mathbf{q}_1$ and $|\mathbf{r}_i| = 1$. Let $\mathbf{q}_3$
sector" of $\mathbf{r}_1$ and $\mathbf{r}_2$. Let $\mathbf{q}_2$ be
d w so that the positive v and w
ly. The picture is shown in Fig-



ısociated with the level surface

If we intersect this figure with the v-w plane, we get Figure 10.

Remark: Note that the choices of orientation of $\mathbf{r}_1$ and $\mathbf{r}_2$ give rise to four possible values of $\mathbf{q}_3$, any two of which are either orthogonal or point in opposite directions. Thus alternate choices may reverse the roles of $\mathbf{q}_2$ and $\mathbf{q}_3$, or reverse their orientations. In the analysis of a shape from the point of view of a given observer, however, there will generally be a "natural" choice of $\mathbf{q}_3$ which is "visible" to the observer.

Let b_i be any C^2 function whose 0 level set is B_i, i.e. B_i: $b_i = 0$. In the u, v, w system $b_i(P_0) = b_i(0, 0, 0) = 0$. Therefore, by Taylor's theorem, we have, near $P_0(0, 0, 0)$,

$$b_i(u, v, w) = \alpha_i u + \beta_i v + \gamma_i w + \epsilon_i(u, v, w),$$

where $\alpha_i = \partial b_i/\partial u(P_0)$, $\beta_i = \partial b_i/\partial v(P_0)$, $\gamma_i = \partial b_i/\partial w(P_0)$. $\epsilon_i(u, v, w)$ is a C^2 function whose order of vanishing at P_0 is greater than 1, (i.e. whose values, and that of its first partial derivatives, is 0 at P_0). Now the intersection locus S is the common set of zeroes $b_1 = b_2 = 0$; the tangent space to S at P_0 is the null space of the Jacobian matrix of b_1, b_2 at P_0, i.e. the null space of

$$\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_1 \end{pmatrix}$$

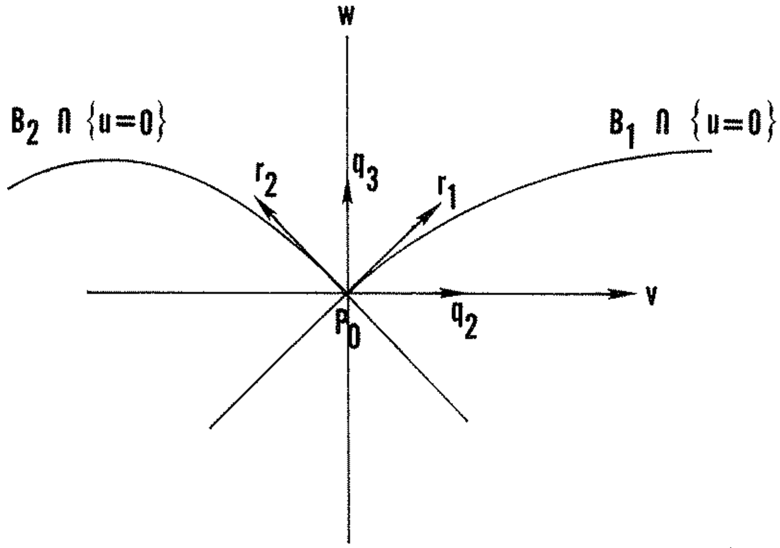


Figure 10. The intersection of Figure 8 with the v, w-plane

in the u, v, w system. However this system was chosen so that the tangent space to S is the u-axis, from which it follows that $\alpha_1 = \alpha_2 = 0$. Moreover, the unit tangent vectors to $B_i \cap \{u = 0\}$, i.e. $\mathbf{r}_1$, $\mathbf{r}_2$ in Figure 10 above, are reflections of each other about the w-axis, and it follows from this that $(\beta_2, \gamma_2) = k(-\beta_1, \gamma_1)$, for some constant k. Thus we may assume

$$b_1 = \beta v + \gamma w + \epsilon_1(u, v, w),$$

$$b_2 = -\beta v + \gamma w + \epsilon_2(u, v, w),$$

for suitable numbers β, γ.

But then, near P_0, $B_1 \cup B_2$ is the level set $b = 0$, where

$$b = -\beta^2 v + \gamma^2 w + \epsilon(u, v, w),$$

(and now ϵ is a function which vanishes at P_0 to order greater than 2).

Now since ϕ and b have the same zero locus, there exists a non-vanishing function h so that $\phi = hb$. At $P_0 = (0, 0, 0)$, $h(u, v, w) = k+$ higher order terms in (u, v, w) where $k \in \mathbf{R}$, $k \neq 0$. Thus, near P_0,

$$\phi(u, v, w) = -k\beta^2 v + k\gamma^2 w + E(u, v, w),$$

where $E(u, v, w)$ is C^2, and vanishes at P_0 to order greater than 2. Letting $l = k\beta^2$, $m = k\gamma^2$, we can summarize as follows:

Proposition 2: Let ϕ be a C^2 function on a domain D in $\mathbf{R}^3$, and suppose M: $\phi = 0$ is the union of two smooth surfaces (the "branches") which intersect in a smooth curve S. Then there exists a right-handed orthogonal coordinate system u, v, w in which P_0 is the origin, and near P_0

$$\phi(u, v, w) = -lv^2 + mw^2 + E(u, v, w),$$

where: (a) l, m are distinct positive numbers such that the tangent planes to the branches are $\sqrt{l}\,v + \sqrt{m}\,w = 0$ and $-\sqrt{l}\,v + \sqrt{m}\,w = 0$ respectively, (b) $E(u, v, w)$ is a C^2 function which vanishes at P_0 to order greater than 2, i.e. its partial derivatives through the second order are zero at P_0, and (c) the u, v, w coordinate system is uniquely determined up to a rotation through a multiple of $\pi/2$ in the $v\,w$ plane (corresponding to the four possible choices of $\mathbf{q}_3$ as in the remark above), and reversal of direction of the w-axis.

We will refer to this expression of ϕ in the u, v, w as the *canonical presentation of ϕ at P_0*; the u, v, w system itself is called the *canonical coordinate system.*

We now want to study the nearby level surfaces $\phi = t$, and especially certain features of their behavior at $t \to 0$. At any point P, we can consider the level set $M(\phi, P)$ of ϕ through P and as $P \to P_0$, $M(\phi, P) \to M(\phi, P_0) =$

M. Let us restrict our atte
smooth; by Sard's theorem (
position 2), it is seen that th
Thus, for $P \notin S$, there is a w
here $N(\phi, P) = \nabla\phi(P)/|\nabla$
ever, a natural question (and
$\lim_{P \to P_0} N(\phi, P)$ exists.

Proposition 3: Let γ b
tangent vector (a, b, c) at
ϕ). Then

$$\lim_{P \to P_0} N(\phi,$$

Thus, the limit depends onl
provided that this direction i

Proof: Assume we have
nate system u, v, w. Let γ
with unit tangent vector
$\gamma(s) = (as + A(s), bs +$
vanish at $s = 0$ to order at
Now $\nabla\phi(u, v, w) = (E_u,$

$$\nabla\phi(\gamma(s)) = [E_u(\gamma(s)), -2lbs$$

$$= s\left[(0, -2lb, 2mc) + \left(\frac{E_u(}{}\right.\right.$$

Let us denote by $\mathbf{F}(s)$
$2mC(s) + E_w(\gamma(s))/s)$. W
and $C(s)$ are actually divis
s, so it is clear that these v
partials of E at $(0, 0, 0)$ va
position 2 above), the parti
we may use

Lemma: Let G be a C^2
G_w all vanish at $(0, 0, 0$
curve γ with $\gamma(0) = 0$.

Proof of Lemma:

$$\lim_{s \to 0} \frac{G(\gamma(s))}{s} = \lim_{s \to 0}$$

was chosen so that the tangent that $\alpha_1 = \alpha_2 = 0$. Moreover, $\mathbf{r}_1$, $\mathbf{r}_2$ in Figure 10 above, are and it follows from this that Thus we may assume

ι, v, w),

u, v, w),

› = 0, where

u, v, w),

o order greater than 2). ıs, there exists a non-vanishing 0), $h(u, v, w) = k+$ higher Thus, near P_0,

+ $E(u, v, w)$,

ɔ order greater than 2. Letting ws: ι domain D in $\mathbf{R}^3$, and suppose s (the "branches") which inter- ight-handed orthogonal coordi- nd near P_0

+ $E(u, v, w)$,

such that the tangent planes to $\overline{l}v + \sqrt{m}w = 0$ respectively, at P_0 to order greater than 2, rder are zero at P_0, and (c) the ned up to a rotation through a ; to the four possible choices of ction of the w-axis. ι, v, w as the *canonical presen-* called the *canonical coordinate*

ıces $\phi = t$, and especially cer- ı point P, we can consider the P_0, $M(\phi, P) \to M(\phi, P_0) =$

M. Let us restrict our attention to those P near S for which $M(\phi, P)$ is smooth; by Sard's theorem (or using the canonical representation of ϕ of Proposition 2), it is seen that there is a neighborhood of S on which this is true. Thus, for $P \notin S$, there is a well defined unit normal $N(\phi, P)$ to $M(\phi, P)$ at P; here $N(\phi, P) = \nabla\phi(P)/|\nabla\phi(P)|$. Now as $P \to P_0$, $|\nabla\phi(P)| \to 0$. However, a natural question (and one which is important for the sequel) is whether $\lim_{P \to P_0} N(\phi, P)$ exists.

Proposition 3: Let γ be a differentiable curve through P_0, with unit tangent vector (a, b, c) at P_0 (in the canonical u, v, w coordinate system for ϕ). Then

$$\lim_{P \to P_0} N(\phi, P) = \frac{(0, -lb, mc)}{|(0, -lb, mc)|}, \quad P \in \gamma.$$

Thus, the limit depends only on the tangent direction of γ at P_0, and it exists provided that this direction is not that of the u-axis, i.e., γ is not tangent to S.

Proof: Assume we have a canonical presentation of ϕ at P_0 with coordinate system u, v, w. Let γ be a differentiable curve through $P_0 = (0, 0, 0)$, with unit tangent vector (a, b, c) at P_0. Thus if s is arclength, then $\gamma(s) = (as + A(s), bs + B(s), cs + C(s))$, where $A(s)$, $B(s)$, and $C(s)$ vanish at $s = 0$ to order at least 2 (i.e. $A(s)$, $B(s)$, $C(s)$ are divisible by s^2). Now $\nabla\phi(u, v, w) = (E_u, -2lv + E_v, 2mv + E_w)$, so that

$$\nabla\phi(\gamma(s)) = [E_u(\gamma(s)), -2lbs - 2lB(s) + E_v(\gamma(s)), 2mcs + 2mC(s) + E_w(\gamma(s))]$$

$$= s\left[(0, -2lb, 2mc) + \left(\frac{E_u(\gamma(s))}{s}, \frac{-2lB(s) + E_v(\gamma(s))}{s}, \frac{-2mC(s) + E_w(\gamma(s))}{s}\right)\right]$$

Let us denote by $\mathbf{F}(s)$ the vector $(E_u(\gamma(s))/s, -2lB(s) + E_v(\gamma(s))/s, 2mC(s) + E_w(\gamma(s))/s)$. We claim that $\mathbf{F}(s)$ vanishes as $s \to 0$, Since $B(s)$ and $C(s)$ are actually divisible by s^2, $B(s)/s$ and $C(s)/s$ are still divisible by s, so it is clear that these vanish. For E_u/s, E_v/s, E_w/s, we note that since the partials of E at $(0, 0, 0)$ vanish through the second order (by part (b) of Proposition 2 above), the partials of E_u, E_v, E_w vanish through the first order, so we may use

Lemma: Let G be a C^2 function around $(0, 0, 0)$ and suppose G, G_u, G_v, G_w all vanish at $(0, 0, 0)$. Then $\lim_{s \to 0} G(\gamma(s))/s = 0$ for any differentiable curve γ with $\gamma(0) = 0$.

Proof of Lemma:

$$\lim_{s \to 0} \frac{G(\gamma(s))}{s} = \lim_{s \to 0} \frac{G(\gamma(s)) - G(\gamma(0))}{s} = \frac{d}{ds} G(\gamma(s))\,|_{s=0}.$$

But this derivative may also be computed in the form

$$\frac{d}{ds}G(\gamma(s)) = (G_u(\gamma(s)),\ G_v(\gamma(s)),\ G_w(\gamma(s)))\gamma'(0),$$

and this is zero by hypothesis.

Returning to the proof of Proposition 3, we have

$$\nabla\phi(\gamma(s)) = s[(0,\ 2lb,\ 2mc) + \mathbf{F}(s)],$$

where $\lim_{s\to 0}\mathbf{F}(s) = (0,\ 0,\ 0)$. Hence

$$\|\nabla\phi(\gamma(s))\| = |s|(\|(0,\ -2lb,\ 2mc)\| + \mathbf{h}(s)),$$

where $\mathbf{h}(s) \to 0$ as $s \to 0$, as is easily seen. Therefore

$$\frac{\nabla\phi(\gamma(s))}{\|\nabla\phi(\gamma(s))\|} = \pm\frac{(0,\ -2lb,\ 2mc) + \mathbf{F}(s)}{\|(0,\ -2lb,\ 2mc)\| + \mathbf{h}(s)},$$

so that

$$\lim_{s\to 0}\frac{\nabla\phi(\gamma(s))}{\|\nabla(\gamma(s))\|} = \pm\frac{(0,\ -2lb,\ 2mc)}{\|(0,\ -2lb,\ 2mc)\|},\ \text{i.e.}$$

$$\lim_{s\to 0}\mathbf{N}(\phi,\ \gamma(s)) = \pm\frac{(0,\ -lb,\ mc)}{\|(0,\ -lb,\ mc)\|}.$$

Note that the sign on the right will be + when $s/\|s\| = 1$, i.e. when $P \to P_0$ from the positive direction with respect to the parametrization of γ. To obtain the result for approach to P_0 along γ from the other direction, simply reverse the orientation of the arc-length parametrization. This completes the proof of proposition 3.

To introduce the basic idea of our main theorem which will be treated in detail in the next section, we now state a prototype of this result, which describes the behavior of the curvature of the level surfaces of ϕ itself at points P approaching P_0.

Proposition 4: With notation and hypotheses as above, let γ be a differentiable curve whose unit tangent vector, say $(a,\ b,\ c)$ in the u, v, w system, is not contained in the tangent planes to the branches B_i of $\phi = 0$ at P_0, and is in the same sector formed by these planes as the positive w-axis. For any $P \notin S$ let $\mu(\phi,\ P)$ denote the mean curvature of the level surface $M(\phi,\ P)$ of ϕ through P.[1] (Thus $\mu(\phi,\ P) = \kappa_1(\phi,\ P) + \kappa_2(\phi,\ P)$). Then

$$\lim_{\substack{P\to P_0\\ P\in\gamma}}\mu(\phi,\ P) = -\infty.$$

[1] It is understood that we are restricting ourselves to a neighborhood of P_0 in which the only non-smooth level surface of ϕ i s $\phi = 0$, i.e. it is only on S that $\nabla\phi$ vanishes.

Proof: By Proposition 1,

$\mu(\phi,\ P) = \frac{\mathrm{tr}H(}{}$

This may be computed in an system for convenience. The tion of P, and

$H(\phi,$

Hence, $\lim_{P\to P_0}\mathrm{tr}H(\phi,\ P) = -$

is the unit vector $(0,\ -lb,\ m$

$\lim_{\substack{P\to P_0\\ P\in\gamma}} N(\phi,\ P)^t H(\phi,\ P)N(\phi,$

$= (l^2b^2 +$

Therefore, (*) the limit of th is

$-l +$

which simplifies to $(-mc^2$ ting this equation to 0, and The set of vectors $(a,\ b,\ c)$ planes $-\sqrt{l}b + \sqrt{m}c = 0$ the branches of $\phi = 0$ at P_0 tors $(a,\ b,\ c)$ which yield the form

$(s)), G_w(\gamma(s)))\gamma'(0),$

/e have

$mc) + \mathbf{F}(s)],$

$b, 2mc) \| + \mathbf{h}(s)),$

Therefore

$\frac{lb, 2mc) + \mathbf{F}(s)}{lb, 2mc) \| + \mathbf{h}(s)},$

$\frac{-2lb, 2mc)}{-2lb, 2mc) \|}$, i.e.

$\frac{), -lb, mc)}{), -lb, mc) \|}.$

hen $s/\|s\| = 1$, i.e. when $P \to P_0$ ie parametrization of γ. To obtain he other direction, simply reverse ition. This completes the proof of

theorem which will be treated in prototype of this result, which the level surfaces of ϕ itself at

potheses as above, let γ be a tor, say (a, b, c) in the u, v, w s to the branches B_i of $\phi = 0$ at ese planes as the positive w-axis. in curvature of the level surface $\kappa_1(\phi, P) + \kappa_2(\phi, P))$. Then

$-\infty.$

o a neighborhood of P_0 in which the only on S that $\nabla\phi$ vanishes.

Proof: By Proposition 1,

$$\mu(\phi, P) = \frac{\mathrm{tr}H(\phi, P) - N(\phi, P)^t H(\phi, P)\mathbf{N}(\phi, P)}{\| \nabla\phi(P) \|}$$

This may be computed in any coordinate system; we will work in the u, v, w system for convenience. Then, $H(\phi, P)$ is a continuous matrix-valued function of P, and

$$H(\phi, P_0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -l & 0 \\ 0 & 0 & m \end{bmatrix}.$$

Hence, $\lim_{P \to P_0} \mathrm{tr}H(\phi, P) = -l + m$. By Proposition 3

$$\lim_{\substack{P \to P_0 \\ P \in \gamma}} N(\phi, P)$$

is the unit vector $(0, -lb, mc)/\sqrt{l^2b^2 + m^2c^2}$. By continuity, we obtain

$$\lim_{\substack{P \to P_0 \\ P \in \gamma}} N(\phi, P)^t H(\phi, P) N(\phi, P)$$

$$= (l^2b^2 + m^2c^2)^{-1}(0, -lb, mc) \begin{bmatrix} 0 & 0 & 0 \\ 0 & -l & 0 \\ 0 & 0 & m \end{bmatrix} \begin{bmatrix} 0 \\ -lb \\ mc \end{bmatrix}$$

$$= \frac{-l^3b^2 + m^3c^2}{l^2b^2 + m^2c^2},$$

Therefore, (*) the limit of the numerator in the expression for $\mu(\phi, P)$ above is

$$-l + m - \left(\frac{-l^3b^2 + m^3c^2}{l^2b^2 + m^2c^2} \right),$$

which simplifies to $(-mc^2 + lb^2)d$ where $d = (l^2b^2 + m^2c^2)^{-1} > 0$. Setting this equation to 0, and solving for b and c, we get $-mc^2 + lb^2 = 0$. The set of vectors (a, b, c) for which this relation holds is the union of the planes $-\sqrt{l}b + \sqrt{m}c = 0$ and $\sqrt{l}b + \sqrt{m}c = 0$, i.e. the tangent planes to the branches of $\phi = 0$ at P_0 as shown in Figure 11. Similarly, the set of vectors (a, b, c) which yield a *negative* numerator in the limit are those

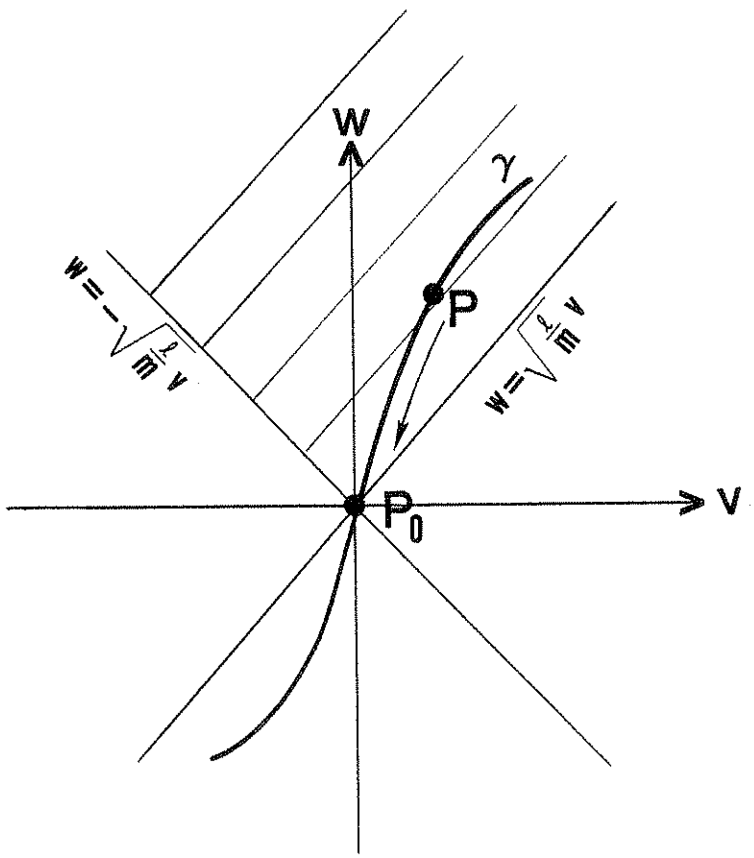


Figure 11. Cross-section through P_0 (in the v, w-plane) of the tangent planes to the branches of M at P_0; the limit of the mean curvatures, $\mu(\psi, P$, as $P \to P_0$ along γ in the shaded sector as shown, is negative because of the orientation of our canonical coordinate system

(a, b, c) for which $|c/b| > \sqrt{l}/\sqrt{m}$, i.e. those which lie on the same sector (formed by the tangent planes to the branches) as the positive w-axis.

Thus let γ be a curve satisfying the hypothesis. As $P \to P_0$ along γ the limit of the numerator in the expression for $\mu(\phi, P)$ is a negative number. Since $\lim_{P \to P_0} \| \nabla\phi \| = 0$, i.e. since the denominator approaches 0 through positive values, we get the result.

We now want to study the question of *uniformity* of approach of $\mu(\phi, P)$ to $-\infty$ as $P \to P_0$, in suitably restricted regions:

Remark: Let $k > \sqrt{l/m}$, and consider the solid cone C in $\mathbf{R}^3$ with vertex P_0, directrix the positive w-axis, and slope k, as shown in Figure 12.

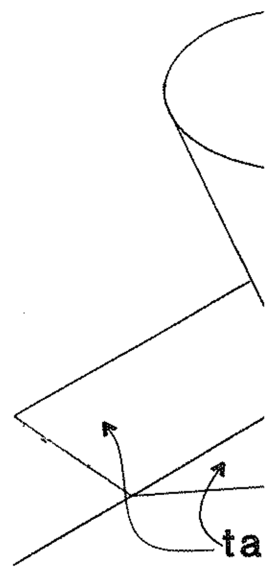
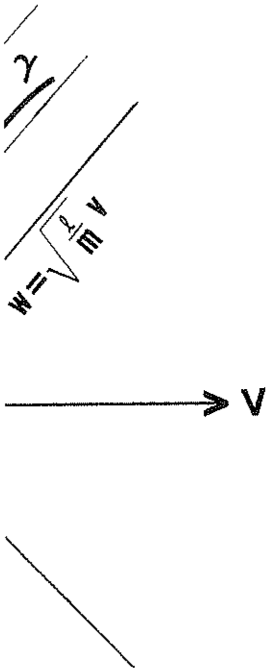


Figure 12. Restricting our atten
figure will enable us to conclude th
through P, goes to $-\infty$ as $P \to P$

Note that, except at P_0, the
$w = \pm\sqrt{l/m}\, v$, i.e. except at
formed by these planes.

Consider the set V of all un
curve with $\gamma'(P_0) = \mathbf{v}$, then th
the expression for $\mu(\phi, P)$ in
$\mathbf{v}$. Moreover, from the explici
proof of Proposition 4, it is cle
compact, there is a maximum
must still be strictly negative f
there must be a vector in V
branches of $\phi = 0$, and this co



ılane) of the tangent planes to the
s, $\mu(\psi, P$, as $P \to P_0$ along γ in
the orientation of our canonical

ɜ which lie on the same sector
s the positive w-axis.
ıesis. As $P \to P_0$ along γ the
(ϕ, P) is a negative number.
ninator approaches 0 through

rmity of approach of $\mu(\phi, P)$
:
olid cone C in $\mathbf{R}^3$ with vertex
shown in Figure 12.

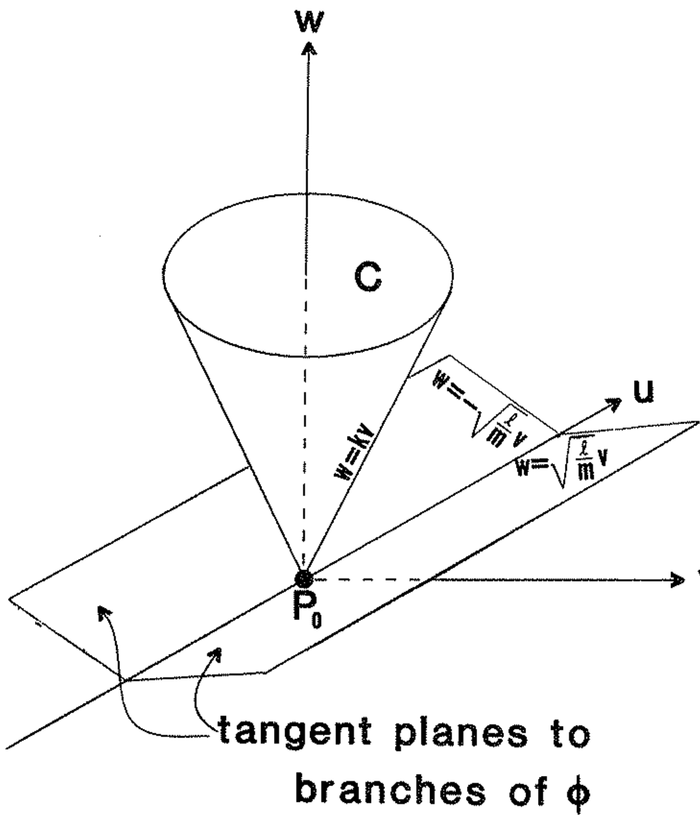


Figure 12. Restricting our attention to points P in cones of the type of C in this figure will enable us to conclude that the mean curvature at P of the level surface of ψ through P, goes to $-\infty$ as $P \to P_0$

Note that, except at P_0, the boundary of this cone does not meet the planes $w = \pm\sqrt{l/m}\,v$, i.e. except at P_0 it is strictly contained in the upper sector formed by these planes.

Consider the set V of all unit vectors lying in C. For each $\mathbf{v} \in V$, if γ is a curve with $\gamma'(P_0) = \mathbf{v}$, then the limit as $P \to P_0$ along γ of the numerator of the expression for $\mu(\phi, P)$ in Proposition 4 is negative, and depends only on $\mathbf{v}$. Moreover, from the explicit expression for this numerator derived in the proof of Proposition 4, it is clear that it depends continuously on $\mathbf{v}$. Since V is compact, there is a maximum value μ for this numerator attained on V, and it must still be strictly negative for otherwise (as in the proof of Proposition 4) there must be a vector in V which lies in the tangent plane to one of the branches of $\phi = 0$, and this contradicts the construction of V.

Now for $P \in C$, consider the line joining P_0 with P, viewed as a curve γ parametrized so that $\gamma(0) = P_0$ and $P \in \gamma$. Let $n(P)$ denote the value of the numerator of $\mu(\phi, P)$, and let $\tilde{n}(P)$ denote its limit as $P \to P_0$ along this γ. Since $\tilde{n}(P)$ depends only on $\gamma'(P_0)$, i.e. in this case on the unit vector in the direction $\overline{P_0P}$, it is clear that $\tilde{n}(P)$ is a continuous function of P and that $\tilde{n}(P) \leqslant \mu < 0$ for all $P \in C$. If $\bar{C}$ denotes the truncation of C at some convenient value of w,[2] then $\bar{C}$ is *compact*, and the function $\tilde{n}(P)$ restricted to $\bar{C}$ is therefore *uniformly* continuous. Moreover, the function $n(P)$ itself is continuous in P, so n is also uniformly continuous on $\bar{C}$.

Consider for the moment points Q on the w-axis; for all these points $\tilde{n}(Q)$ is the same (and in fact equals $-l$). There exists an $\epsilon_1 > 0$ such that if $|Q - P_0| < \epsilon_1$, then $|n(Q) - \tilde{n}(Q)| < |\mu|/4$. By the uniform continuity of n, there exists ϵ_2 so that for all $P, P' \in \bar{C}$, $|P - P'| < \epsilon_2 \Rightarrow |n(P) - n(P')| < |\mu|/4$. Let ϵ_0 be sufficiently small so that $\epsilon_0 < \epsilon_1$, and moreover that the cross-section of $\bar{C}$ slant height ϵ_0 has radius less than ϵ_2 (see Figure 13). Note that then if $P \in \bar{C}$ with $|P - P_0| < \epsilon_0$, then if Q denotes the point on the w-axis with the same w-coordinate as P, we have both $|P - Q| < \epsilon_2$ and $|Q - P_0| < \epsilon_1$.

It follows that

$$|n(P) - \tilde{n}(Q)| \leqslant |n(P) - n(Q)| + |n(Q) - \tilde{n}(Q)| \leqslant \frac{|\mu|}{4} + \frac{|\mu|}{4} = \frac{|\mu|}{2}.$$

Since $\tilde{n}(Q) \leqslant \mu$, we find: There exists ϵ_0 so that $P \in \bar{C}$, $|P - P_0| \leqslant \epsilon_0 \Rightarrow n(P) \leqslant \mu/2$ (μ a fixed negative number).

Now since ϕ is C^2, $\| \phi(P) \| \to 0$ uniformly on $\bar{C}$ as $P \to P_0$. It follows that: $\mu(\phi, P) \to -\infty$ uniformly on $\bar{C}$ as $P \to P_0$. We may summarize this in the following result which may now be stated in a coordinate free form:

Theorem 5. Let ϕ be a C^2 function on a domain D, in $\mathbf{R}^3$. Suppose the level set $\phi = 0$ consists of two smooth surfaces B_1 and B_2 which intersect in a smooth curve S; suppose moreover that all other level sets of ϕ in D are smooth. Let $P_0 \in S$, and let L be a line through P_0, perpendicular to S, which bisects the tangent planes to the branches B_i at P_0. Let C be any solid cone in D with vertex P_0 and directrix L, which does not touch the tangent planes to the branches except at P_0 itself. Then

$$\lim_{\substack{P \to P_0 \\ P \in C}} \mu(\phi, P) = -\infty.$$

where $\mu(\phi, P)$ is the mean curvature at P of the level surface of ϕ through P. Moreover this limit is approached *uniformly* on C.

[2] All that is required is that $\bar{C}$ be contained in the domain of definition of ϕ, and that $|\nabla\phi(P)| \neq 0$ for all $P \in \bar{C}$ except $P = P_0$.

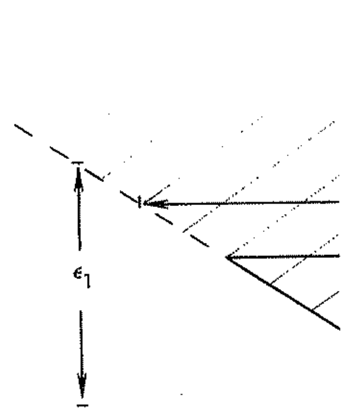


Figure 13. An illustration of tl
mean curvatures, $\mu(\psi, P)$, as P

The main theorem. In g
level set $\phi = 0$ has singula
singular along the curve S), th
smooth. Thus we can think of
family of smooth approxima
some suitable geometric sens
constructing such "smoothing
will result in functions with s

In addition to the $\phi = t$ s
specific mention at this point.
the point here is to motivate
will be stated. First, there i
define

$$f_t(x) =$$

The $f_t(x)$ can be viewed a
small t, $u \in R$ the level sets
Note that $\lim_{t \to 0} f_t = \phi$, and if
differentiably on t and x.

P_0 with P, viewed as a curve γ
Let $n(P)$ denote the value of the
ts limit as $P \to P_0$ along this γ.
iis case on the unit vector in the
ntinuous function of P and that
he truncation of C at some con-
he function $\tilde{n}(P)$ restricted to $\bar{C}$
the function $n(P)$ itself is con-
s on $\bar{C}$.
w-axis; for all these points $\tilde{n}(Q)$
exists an $\epsilon_1 > 0$ such that if
$\mu|/4$. By the uniform continuity
all P, $P' \in \bar{C}$, $|P - P'|$
$_0$ be sufficiently small so that
of $\bar{C}$ slant height ϵ_0 has radius
if $P \in \bar{C}$ with $|P - P_0| < \epsilon_0$,
the same w-coordinate as P, we
ϵ_1.

$$Q)| \leqslant \frac{|\mu|}{4} + \frac{|\mu|}{4} = \frac{|\mu|}{2}.$$

exists ϵ_0 so that $P \in \bar{C}$,
negative number).
ly on $\bar{C}$ as $P \to P_0$. It follows
P_0. We may summarize this in
in a coordinate free form:
domain D, in $\mathbf{R}^3$. Suppose the
ces B_1 and B_2 which intersect in
other level sets of ϕ in D are
h P_0, perpendicular to S, which
t P_0. Let C be any solid cone in
not touch the tangent planes to

$-\infty$.

the level surface of ϕ through P.
n C.

le domain of definition of ϕ, and that

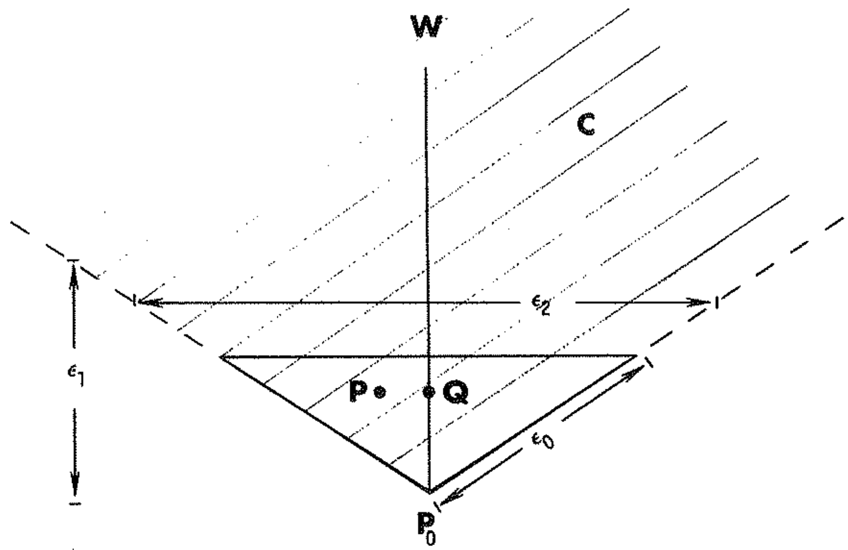


Figure 13. An illustration of the argument arising in the study of the limit of the mean curvatures, $\mu(\psi, P)$, as $P \to P_0$ in C

The main theorem. In general, if ϕ is a differentiable function whose level set $\phi = 0$ has singularities (for example as above, where $\phi = 0$ is singular along the curve S), the nearby level sets $\phi = t$ ($t \in R$, t small) will be smooth. Thus we can think of the family of surfaces M_t: $\phi = t$ as a canonical family of smooth approximations to M_0: $\phi = 0$ which converge to M_0 in some suitable geometric sense as $t \to 0$. However there are many ways of constructing such "smoothing families," and in fact generic perturbations of ϕ will result in functions with smooth level surfaces near $\phi = 0$.

In addition to the $\phi = t$ smoothing, we single out two other examples for specific mention at this point. We will return to them in more detail later (); the point here is to motivate the level of generality at which the Theorem 6 will be stated. First, there is "Gaussian smoothing," where for $t > 0$ we define

$$f_t(x) = \frac{1}{\sqrt{2\pi t}} \int_D e^{\frac{-|x-v|^2}{2t^2}} \phi(v)dv.$$

The $f_t(x)$ can be viewed as "smoothings of the function ϕ" itself, and for small t, $u \in R$ the level sets $f_t = u$ are smoothings of the surface $\phi = 0$. Note that $\lim_{t \to 0} f_t = \phi$, and if we define $f_0 = \phi$ the function $f_t(x)$ depends differentiably on t and x.

As a second example, start with a sequence of lattices (i.e. discrete subsets) L_i of D, which get arbitrarily dense in D as $i \to \infty$. We can use each lattice as a set of control points to construct a suitable bicubic spline f_i, so that $\lim_{i \to \infty} f_i = \phi$ and the level sets of the f_i near $\phi = 0$ are smooth. This example seems fundamentally different than that of the Gaussian smoothing; for one thing the "parameter" i is discrete, and there is no "natural" continuous parameter as in the Gaussian case. Note however that any "continuous parameter" smoothing f_t of ϕ can be represented as a suitable sequence f_i converging to ϕ, without excessive information loss, at least at ϕ itself. For this reason we will consider a sequence of functions f_i (with smooth level surfaces) which converge to ϕ in C^2, as the most general practical formulation of a "smoothing family."

Theorem 6, which we prove in this section, states that when we smooth a transversal intersection in this sense, negative curvature at points of the smooth surfaces increases without bound the closer we get to the singular level set $\phi = 0$. Thus high negative curvature is seen to be the "stable form" of transversal intersections, in the sense that slight perturbations of the singularity yield smooth surfaces with arbitrarily high curvature. Indeed, it is impossible to actually detect an intersection in practice (if we have access to only one side of the surface); the most we can do is to measure curvature up to the order of $1/\sigma$, where σ is the limit of the resolving power of our measurement system.

Theorem 6: Let D be a domain in R^3, and let $\phi \in C^2(D)$ be a function whose level set $\phi = 0$ is the union of two smooth branches which intersect transversally in a smooth curve S. Suppose $\{f_i\}$ is a sequence of functions which converge in $C^2(D)$ to ϕ, and all the level sets of the f_i through any point $P \in D - S$ are smooth. Let γ be a differentiable curve in D which intersects $\phi = 0$ at a single point $P_0 \in S$, and which is not tangent to either of the branches of $\phi = 0$. Then

$$\lim_{\substack{i \to \infty \\ \delta \to 0}} \inf\{\mu(f_i, P) \mid P \in \gamma, \mid P - P_0 \mid \leqslant \delta\} = -\infty$$

where $\mu(f_i, P)$ denotes the mean curvature at P of $M(f_i, P)$, i.e. the level set of f_i through P.

Remark: If we let $f_i = \phi$ for each i in Theorem 6, we get a form of Theorem 5.

Proof of Theorem 6: As in Proposition 1 of §3.1, for any $P \in D - S$ we may write

$$\mu(f_i, P) = \frac{\operatorname{tr} H(f_i, P) - \mathrm{N}(f_i, P)^t H(f_i, P) \mathrm{N}(f_i, P)}{\mid \nabla f_i(P) \mid}$$

where $H(f, P)$ denotes the Hessi
normal to the level surface $M(f,$
the direction of $\nabla f(P)$. To sim
$\langle\, ,\rangle_{i,P}$, and its trace by $t_i(P)$. I
$\langle\, ,\rangle_P$ and $t(P)$ respectively. Mo
$\mathrm{N}_i(P)$, and $\mathrm{N}(\phi, P)$ by $\mathrm{N}(P)$.

Note that $\mathrm{N}(P_0)$ is not a p
$\lim \mathrm{N}(P)$ exists as $P \to P_0$ along
branches of $\phi = 0$ and in partic
the notation $\mathrm{N}(P_0)$ to denote th
throughout the discussion there is

With our new notation, we hav

$$\mu(f_i, P) = \frac{t_i(}{}$$

We add and subtract the qu
numerator to obtain:

$$\mu(f_i, P) =$$

$$(t(P_0) - \langle \mathrm{N}(P_0), \mathrm{N}(P_0) \rangle_{P_0}) + ($$

Now the term on the left in th
the numerator of the expression
noting that our γ satisfies the h
orientation of the axes. Therefor
term is a fixed negative number k

We now expand the term on t
tracting $\langle \mathrm{N}(P_0), \mathrm{N}(P_0) \rangle_{i,P_0}$, to c

$$\mu(f_i, P) = k / |\nabla f_i(P)|$$
$$+ \ (t_i(P) - t(P_0$$
$$- \ (\langle \mathrm{N}_i(P), \mathrm{N}_i(P$$
$$+ \ (\langle \mathrm{N}(P_0), \mathrm{N}(P$$

In order to prove the theorem
positive integer. Then there e

ce of lattices (i.e. discrete sub-
as $i \to \infty$. We can use each lat-
uitable bicubic spline f_i, so that
$= 0$ are smooth. This example

ie Gaussian smoothing; for one
ere is no "natural" continuous
ver that any "continuous param-
a suitable sequence f_i converg-
, at least at ϕ itself. For this
ons f_i (with smooth level sur-
general practical formulation of

, states that when we smooth a
ve curvature at points of the
closer we get to the singular
is seen to be the "stable form"
light perturbations of the singu-
high curvature. Indeed, it is
practice (if we have access to
do is to measure curvature up
resolving power of our meas-

id let $\phi \in C^2(D)$ be a function
nooth branches which intersect
$f_i\}$ is a sequence of functions
evel sets of the f_i through any
ntiable curve in D which inter-
h is not tangent to either of the

$_0| \leqslant \delta\} = -\infty$

of $M(f_i, P)$, i.e. the level set

Theorem 6, we get a form of

f §3.1, for any $P \in D - S$ we

$)^t H(f_i, P) \mathrm{N}(f_i, P)$
$)|$

where $H(f, P)$ denotes the Hessian form of f at P and $\mathrm{N}(f, P)$ is the unit normal to the level surface $M(f, P)$ at P, i.e. $\mathrm{N}(f, P)$ is the unit vector in the direction of $\nabla f(P)$. To simply notation we will denote $H(f_i, P)$ by $\langle\, ,\rangle_{i,P}$, and its trace by $t_i(P)$. $H(\phi, P)$ and $\mathrm{tr}H(\phi, P)$ will be denoted by $\langle\, ,\rangle_P$ and $t(P)$ respectively. Moreover we will denote $\mathrm{N}(f_i, P)$ simply by $\mathrm{N}_i(P)$, and $\mathrm{N}(\phi, P)$ by $\mathrm{N}(P)$.

Note that $\mathrm{N}(P_0)$ is not a priori defined. However by Proposition 3, $\lim \mathrm{N}(P)$ exists as $P \to P_0$ along γ, since by hypothesis γ is not tangent to the branches of $\phi = 0$ and in particular not tangent to S. We will therefore use the notation $\mathrm{N}(P_0)$ to denote this limit for our given γ; since γ is fixed throughout the discussion there is no ambiguity.

With our new notation, we have

$$\mu(f_i, P) = \frac{t_i(P) - \langle \mathrm{N}_i(P), \mathrm{N}_i(P)\rangle_{i,P}}{|\nabla f_i(P)|}$$

We add and subtract the quantity $t(P_0) - \langle \mathrm{N}(P_0), \mathrm{N}(P_0)\rangle_{P_0}$ in the numerator to obtain:

$$\mu(f_i, P) = \frac{(t(P_0) - \langle \mathrm{N}(P_0), \mathrm{N}(P_0)\rangle_{P_0}) + (t_i(P) - t(P_0)) - (\langle \mathrm{N}_i(P), \mathrm{N}_i(P)\rangle_{i,P} - \langle \mathrm{N}(P_0), \mathrm{N}(P_0)\rangle_{P_0})}{|\nabla f_i(P)|}$$

Now the term on the left in the numerator is the limit as $P \to P_0$ along γ of the numerator of the expression for $\mu(\phi, 0)$ as in the proof of Proposition 4, noting that our γ satisfies the hypotheses of that proposition for a suitable orientation of the axes. Therefore, by (*) in the proof of Proposition 4, this term is a fixed negative number k.

We now expand the term on the right in the numerator by adding and subtracting $\langle \mathrm{N}(P_0), \mathrm{N}(P_{0)}\rangle_{i,P_0}$, to obtain:

$$\begin{aligned}
\mu(f_i, P) &= k/|\nabla f_i(P)| \\
&\quad + (t_i(P) - t(P_0))/|\nabla f_i(P)| \qquad (6.1)\\
&\quad - (\langle \mathrm{N}_i(P), \mathrm{N}_i(P)\rangle_{i,P} - \langle \mathrm{N}(P_0), \mathrm{N}(P_0)\rangle_{i,P})/|\nabla f_i(P)| \\
&\quad + (\langle \mathrm{N}(P_0), \mathrm{N}(P_0)\rangle_{P_0} - \langle \mathrm{N}(P_0), \mathrm{N}(P_0)\rangle_{i,P})/|\nabla f_i(P)|.
\end{aligned}$$

In order to prove the theorem, we must show the following: let L be any positive integer. Then there exist n, δ such that there is a $P \in \gamma$ with

$|P - P_0| \leqslant \delta$ and $\mu(f_i, P) < -L$ for all $i > n$. In fact we will prove the stronger assertion

(6.2) Given L, there exist n, δ such that for $i > n$ and *all* P satisfying $\delta/2 \leqslant |P - P_0| \leqslant \delta$, $\mu(f_i, P) < -L$.

We will denote

$$\bar{1} = t_i(P) - t(P_0),$$

$$\bar{2} = \langle \mathbf{N}_i(P), \mathbf{N}_i(P)\rangle_{i,P} - \langle \mathbf{N}(P_0), \mathbf{N}(P_0)\rangle_{i,P},$$

$$\bar{3} = \langle \mathbf{N}(P_0), \mathbf{N}(P_0)\rangle_{P_0} - \langle \mathbf{N}(P_0), \mathbf{N}(P_0)\rangle_{i,P}.$$

With this notation, (6.1) becomes

$$\mu(f_i, P) = \frac{k + \bar{1} - \bar{2} + \bar{3}}{|\nabla f_i(P)|}.$$

Recalling that k is a *fixed negative number* (depending only on γ), to prove (6.2) we will find n, δ such that for $i > n$ and $\delta/2 \leqslant |P - P_0| \leqslant \delta$, we have $|\nabla f_i(P)| < |3k/2L|$, $|\bar{1}| < |k|/6$, $|\bar{2}| < |k|/6$, $|\bar{3}| < |k|/6$. This will give the result, for then

$$\mu(f_i, P) \leqslant \frac{k + |\bar{1}| + |\bar{2}| + |\bar{3}|}{|\nabla f_i(P)|}$$

$$\leqslant \frac{k + |k/6| + |k/6| + |k/6|}{|\nabla f_i(P)|}$$

$$= \frac{3k/2}{|\nabla f_i(P)|} \leqslant \frac{3k/2}{|3k/2L|} = -L.$$

Now, since $f_i \to \phi$ in C^2, it follows that $|\nabla f_i| \to |\nabla\phi|$ *uniformly* on compact subsets of γ, and similarly $\langle\, ,\rangle_{i,P} \to \langle\, ,\rangle_P$ uniformly on compact subsets of γ (for this purpose we may identify $\langle\, ,\rangle$ etc. with the appropriate Hessian matrix of second partial derivatives.)

Let us first choose δ_1 so that if $|P - P_0| \leqslant \delta_1$, $|\nabla\phi(P)| < 3|k|/4L$; we can do this since $|\nabla\phi(P_0)| = 0$ and $\nabla\phi$ is continuous. Next, choose n_1 so that, in view of the uniform convergence of $\nabla f_i(P)$ to $\nabla\phi(P)$ on the compact subset of γ defined by $|P - P_0| \leqslant \delta_1$, $|\nabla f_i(P) - \nabla\phi(P)| \leqslant 3|k|/4L$ for P in this subset. It follows that for $i > n_1$ and $|P - P_0| \leqslant \delta_1$, $|\nabla f_i(P)| < 3|k|/2L$.

Now we look at the term $\bar{1}$. We can write it $(t_i(P) - t(P)) + (t(P) - t(P_0))$. Recall that $t_i(p)$ is the trace of $\langle\, ,\rangle_{i,P}$ and in particular it is a sum of second partial derivatives. Since the second

$> n$. In fact we will prove the

or $i > n$ and *all* P satisfying

$P_0), \mathbf{N}(P_0)\rangle_{i,P}$,

$(P_0), \mathbf{N}(P_0)\rangle_{i,P}$.

$\overline{2} + \overline{3}$

$P)|$

lepending only on γ), to prove
ind $\delta/2 \leqslant |P - P_0| \leqslant \delta$, we
$| < |k|/6$, $|\overline{3}| < |k|/6$. This

$|\overline{2}| + |\overline{3}|$

$P)|$

$|k/6| + |k/6|$

$f_i(P)|$

$\leqslant \frac{3k/2}{|3k/2L|} = -L.$

$\nabla f_i| \to |\nabla\phi|$ *uniformly* on com-
$_P$ uniformly on compact subsets
tc. with the appropriate Hessian

$| \leqslant \delta_1$, $|\nabla\phi(P)| < 3|k|/4L$;
is continuous. Next, choose n_1
f $\nabla f_i(P)$ to $\nabla\phi(P)$ on the com-
, $|\nabla f_i(P) - \nabla\phi(P)| \leqslant 3|k|/$
$i > n_1$ and $|P - P_0| \leqslant \delta_1$,

$\overline{1}$. We can write it
hat $t_i(p)$ is the trace of $\langle\, ,\rangle_{i,P}$
al derivatives. Since the second

derivatives of ϕ are continuous, we can find a δ_2 so that if $|P - P_0| \leqslant \delta_2$, $|t(P) - t(P_0)| < |k|/12$. Then, since the second partial derivatives of the f_i converge uniformly to those of ϕ on the compact subset of γ defined by $|P - P_0| \leqslant \delta_2$, we can find n_2 so that if $i > n_2$, $|t_i(P) - t(P)| < |k|/12$ for P in this subset. It follows that for $i > n_2$ and $|P - P_0| < \delta_2$, $\overline{1} < |k|/6$.

We consider the term $\overline{3}$. Let $\mathbf{B}$ denote the unit vector $\mathbf{N}(P_0)$; it depends only on γ so is fixed throughout the discussion. Thus $\overline{3}$ is $\langle \mathbf{B}, \mathbf{B}\rangle_{P_0} - \langle \mathbf{B}, \mathbf{B}\rangle_{i,P}$. We can write this $(\langle \mathbf{B}, \mathbf{B}\rangle_{P_0} - \langle \mathbf{B}, \mathbf{B}\rangle_P) + (\langle \mathbf{B}, \mathbf{B}\rangle_P - \langle \mathbf{B}, \mathbf{B}\rangle_{(i,P)})$. Choose δ_3 (by the continuity of the second partial derivatives of ϕ) so that if $|P - P_0| \leqslant \delta_3$, $|\langle \mathbf{B}, \mathbf{B}\rangle_{P_0} - \langle \mathbf{B}, \mathbf{B}\rangle_P| \leqslant |k|/12$. Then choose n_3 (by the uniform convergence of $\langle\, ,\rangle_{i,P}$ to $\langle\, ,\rangle_P$ on the compact subset of γ for which $|P - P_0| \leqslant \delta_3$) so that if $i > n_3$, $|\langle \mathbf{B}, \mathbf{B}\rangle_P - \langle \mathbf{B}, \mathbf{B}\rangle_{i,P}| < |k|/12$. It follows that there exist n_3, δ_3 so that if $i > n_3$ and $|P - P_0| \leqslant \delta_3$, $\overline{3} < |k|/6$.

Now, to study the term **2** we first write it in the form $\overline{2} = \langle \mathbf{N}_i(P) + \mathbf{N}(P_0), \mathbf{N}_i(P) - \mathbf{N}(P_0)\rangle_{i,P}$. Reverting back to matrix notation, let H denote the Hessian matrix of f_i at P in some coordinate system, and let $N_i(P)$ and $N(P_0)$ be written as vectors in the same system. Then

$$\overline{2} = (\mathbf{N}_i(P) + \mathbf{N}(P_0)\cdot H(\mathbf{N}_i(P) - \mathbf{N}(P_0))$$

(where the dot denotes ordinary dot product in the given coordinate system, and H is operating as a matrix on the vector $\mathbf{N}_i(P) - \mathbf{N}(P_0)$). Hence

$$\begin{aligned}\overline{2} &\leqslant |\mathbf{N}_i(P) + \mathbf{N}(P_0)|\,|H(\mathbf{N}_i(P) - \mathbf{N}(P_0))|,\\ &\leqslant 2|H(\mathbf{N}_i(P) - \mathbf{N}(P_0))|,\end{aligned}$$

since $\mathbf{N}_i(P)$, $\mathbf{N}(P_0)$ are both unit vectors.

It is a well known and easily verified fact that if H is any matrix, and l denotes the sum of the lengths of the columns of H, then for any vector $\mathbf{v}$, $|H(\mathbf{v})| \leqslant l|\mathbf{v}|$. Let $\bar{l}$ denote an upper bound of the values of l obtained for the Hessian matrices $H = H(f_i, P)$ for $P \in \gamma$, $|P - P_0| \leqslant \delta_3$ and $i = 1, 2, \cdots$; to see that this bound exists, observe that there is a similar bound for the Hessian of ϕ at points P in this compact set, and then use the uniform convergence of the Hessians of the f_i to those of ϕ on the set. Thus, if $|P - P_0| \leqslant \delta_3$,

$$\begin{aligned}|\overline{2}| &\leqslant 2\bar{l}|\mathbf{N}_i(P) - \mathbf{N}(P_0)|,\\ &\leqslant 2\bar{l}(|\mathbf{N}_i(P) - \mathbf{N}(P)| + |\mathbf{N}(P) - \mathbf{N}(P_0)|).\end{aligned}$$

Now $\mathbf{N}(P)$ converges to $\mathbf{N}(P_0)$ as $P \to P_0$ along γ. Therefore, choose a δ_4 which is $\leqslant \delta_3$, and which also has the property that if

$|P - P_0| \leqslant \delta_4 |\mathbf{N}(P) - \mathbf{N}(P_0)| \leqslant |k|/(24\bar{l})$. Observe that away from S (i.e. away from P_0 in our situation we are just working on γ), $\mathbf{N}(P)$ is a well-defined vector-valued function of P, expressible in terms of the derivatives of ϕ. Thus on any compact set not meeting S, $\mathbf{N}_i(P)$ converges uniformly to $\mathbf{N}(P)$. In particular, there exists n_4 so that if $i > n_4$ and $\delta_4/2 \leqslant |P - P_0| \leqslant \delta_4$, then $|\mathbf{N}_i(P) - \mathbf{N}(P)| < |k|/(24\bar{l})$. It follows that there exist n_4 and δ_4 so that if $i > n_4$ and $\delta_4/2 \leqslant |P - P_0| \leqslant \delta_4$, $|\bar{2}| < |k|/6$.

It is now clear that if we let $\delta = \inf(\delta_1, \delta_2, \delta_3, \delta_4)$ and $n = \sup(n_1, n_2, n_3, n_4)$ the conclusion of 6.2 is valid. This completes the proof of Theorem 6.

Remark. Suppose that we have a smoothing family $f_i \rightarrow \phi$ as in Theorem 6. We would like to conclude that for sufficiently large i, the level surfaces of the f_i close to $\phi = 0$ contain contours which are part boundaries in the sense of the Negative Minima Partitioning Rule of §2 above, and moreover that these contours converse to S in some reasonable way. Theorem 6 makes this very plausible, but does not go all the way to give a proof. The essential difficulties here may be understood by considerting the case of the level surface smoothing of ϕ, i.e., the case where all the f_i are just ϕ itself. As we have seen, these level surfaces have points of arbitrarily high negative curvature close to S. The problem lies in the possibility that the relevant lines of curvature on the given level surfaces near S get "trapped" in neighborhoods of S. In this way the line of curvature might approach S so that the points on it could have ever increasing negative curvature, i.e., no points is an extremum. Since the boundary contours by definition consist of points which are extrema of curvature on their corresponding lines of curvature, there would be no such contour in this case. Happily this kind of pathology can be ruled out; the methods are beyond the scope of this paper and will be published elsewhere.

4. EXAMPLES OF PARTITIONS

In the previous section we proved that smoothing a transversal intersection leads to large curvature (negative for solid union, positive for solid subtraction) regardless of how one smooths. In this section we determine analytically the negative minima partitioning contours on several classes of surfaces. This allows a more rigorous understanding of the rule and the boundaries it defines. In particular, this section illustrates that *the negative minima rule is a 3-D definition of part boundary, not a 2-D rule of thumb for finding 2-D parts* (such as the "matched concavities heuristic,"—see Brady & Asada, 1984, for a description, and critique, of the matched concavities heuristic).

Decomposition of developal

Developable surfaces are a
generated by a one param
parameter family of lines {
the parameter $u^1 \in (a, b) \subset$
$\mathbf{v}(u^1) \neq 0$, such that both
each $u^1 \in (a, b)$, the line $L($
$\mathbf{v}(u^1)$ is called the line of th

Given a one parameter fa
surface is given by the parar

$$\mathbf{x}(u^1, u^2) = \mathbf{p}(u^1)$$

The curve $\mathbf{p}(u^1)$ is called
the surface $\mathbf{x}$. In what follo
is arc length along $\mathbf{p}$ and th
lowing notation: If w is a v
u^1, u^2, then w_i will denote
$\mathbf{x}_2 = \partial\mathbf{x}/\partial u^2$ etc. A ruled s
product $(\mathbf{v}, \mathbf{v}_1, \mathbf{p}_1) = 0$ ev
$\mathbf{p}_1$ all lie in a single plane.

In the next two subsectio
by the minima rule for two
cylinder and cone.

Cylinders

A cylinder is a developable
plane and whose rulings, $\mathbf{v}($
ing that $\mathbf{v}_1 = 0$. For
$\mathbf{x}_1 = \mathbf{p}_1 + u^2\mathbf{v}_1 = \mathbf{p}_1$ (sinc

$$g_{ij}$$

The surface normal is

$$\mathbf{N} = \frac{\mathbf{x}_1}{|\mathbf{x}_1}$$

The second fundamental

. Observe that away from S
st working on γ), $N(P)$ is a
ssible in terms of the deriva-
:ing S, $N_i(P)$ converges uni-
$_4$ so that if $i > n_4$ and $\delta_4/$
$< |k|/(24\bar{l})$. It follows that
nd $\delta_4/2 \leqslant |P - P_0| \leqslant \delta_4$,

$\inf(\delta_1, \delta_2, \delta_3, \delta_4)$ and $n =$
ılid. This completes the proof

family $f_i \to \phi$ as in Theorem
y large i, the level surfaces of
e part boundaries in the sense
§2 above, and moreover that
› way. Theorem 6 makes this
› give a proof. The essential
ting the case of the level sur-
ıe f_i are just ϕ itself. As we
bitrarily high negative curva-
ility that the relevant lines of
t "trapped" in neighborhoods
proach S so that the points on
ature, i.e., no points is an
nition consist of points which
ing lines of curvature, there
this kind of pathology can be
f this paper and will be pub-

ing a transversal intersection
›n, positive for solid subtrac-
:ion we determine analytically
/eral classes of surfaces. This
rule and the boundaries it
the negative minima rule is a
of thumb for finding 2-D parts
ee Brady & Asada, 1984, for
vities heuristic).

Decomposition of developable surfaces

Developable surfaces are a special case of ruled surfaces, surfaces that are generated by a one parameter family of lines (Do Carmo, 1976). A one parameter family of lines $\{\mathbf{p}(u^1), \mathbf{v}(u^1)\}$ is a correspondence that assigns to the parameter $u^1 \in (a, b) \subset \mathbf{R}$ a point $\mathbf{p}(u^1) \in \mathbf{R}^3$ and a vector $\mathbf{v}(u^1) \in \mathbf{R}^3$, $\mathbf{v}(u^1) \neq 0$, such that both $\mathbf{p}(u^1)$ and $\mathbf{v}(u^1)$ depend differentiably on u^1. For each $u^1 \in (a, b)$, the line $L(u^1)$ which passes through $\mathbf{p}(u^1)$ and is parallel to $\mathbf{v}(u^1)$ is called the line of the family at u^1.

Given a one parameter family of lines $\{\mathbf{p}(u^1), \mathbf{v}(u^1)\}$, the associated ruled surface is given by the parameterization

$$\mathbf{x}(u^1, u^2) = \mathbf{p}(u^1) + u^2\mathbf{v}(u^1), \quad u^1 \in (a, b) \subset \mathbf{R}, \quad u^2 \in \mathbf{R}.$$

The curve $\mathbf{p}(u^1)$ is called a *directrix*, and the lines are called the *rulings* of the surface $\mathbf{x}$. In what follows we assume, without loss of generality, that u^1 is arc length along $\mathbf{p}$ and that $|\mathbf{v}(u^1)| = 1$. Moreover we will adopt the following notation: If w is a vector or scalar valued function of the parameters u^1, u^2, then w_i will denote $\partial w/\partial u^i$. Thus $\mathbf{v} = \partial\mathbf{v}/\partial u^1$, $\mathbf{p}_{11} = \partial^2\mathbf{p}/(\partial u^1)^2$, $\mathbf{x}_2 = \partial\mathbf{x}/\partial u^2$ etc. A ruled surface is said to be *developable* if the scalar triple product $(\mathbf{v}, \mathbf{v}_1, \mathbf{p}_1) = 0$ everywhere on the surface, implying that $\mathbf{v}$, $\mathbf{v}_1$, and $\mathbf{p}_1$ all lie in a single plane.

In the next two subsections we determine the partitioning contours defined by the minima rule for two nonexhaustive cases of developable surfaces—the cylinder and cone.

Cylinders

A cylinder is a developable surface whose directrix, $\mathbf{p}$, lies entirely in one plane and whose rulings, $\mathbf{v}(u^1)$, are parallel to a fixed direction in $\mathbf{R}^3$, implying that $\mathbf{v}_1 = 0$. For a cylinder the first partial derivatives are $\mathbf{x}_1 = \mathbf{p}_1 + u^2\mathbf{v}_1 = \mathbf{p}_1$ (since $\mathbf{v}_1 = 0$) and $\mathbf{x}_2 = \mathbf{v}$. The metric tensor is then

$$g_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The surface normal is

$$\mathbf{N} = \frac{\mathbf{x}_1 \times \mathbf{x}_2}{|\mathbf{x}_1 \times \mathbf{x}_2|} = \frac{\mathbf{p}_1 \times \mathbf{v}}{|\mathbf{p}_1 \times \mathbf{v}|} = \mathbf{p}_1 \times \mathbf{v}$$

The second fundamental coefficients are

$$b_{ij} = \mathbf{x}_{ij}\cdot\mathbf{N} = \begin{pmatrix} (\mathbf{p}_{11}, \mathbf{p}_1, \mathbf{v}) & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} |\mathbf{p}_{11}| & 0 \\ 0 & 0 \end{pmatrix}$$

Since $g_{12} = b_{12} = 0$ the principal curvatures on a cylinder are

$$\kappa_1 = b_{11}/g_{11} = |\mathbf{p}_{11}|$$

$$\kappa_2 = b_{22}/g_{22} = 0$$

The expression for κ_1 is the magnitude of the second derivative of $\mathbf{p}$ with respect to arc length (with sign determined by the orientation of the field of surface normals) which is simply the curvature along the directrix $\mathbf{p}$. The directrix and its translations are, in fact, one set of lines of curvature and the rulings the other set, since g_{12} and b_{12} are zero. As expected, the curvature along the rulings, κ_2, is zero. Consequently no partitioning contours arise from the rulings (since there are no extrema of the principal curvature κ_2). Only the minima of κ_1 along the directrix and its translations are used for defining part boundaries.

Figure 2, as discussed in the introduction, shows a cylinder and its partitioning contours (dotted lines) for one of the orientations of the field of surface normals. The partitioning contours break the cylinder into parts that seem natural enough. If one inverts the figure one will experience a figure-ground reversal, causing the bumps of the surface to become dips and vice-versa. Notice that when the figure and ground reverse the perceived partitioning lines shift away from the indicated dotted lines and to the lines that were previously positive maxima of κ_1. This occurs because the figure-ground reversal is associated with a reversal in the orientation of the field of surface normals and, hence, in the sign of κ_1 everywhere on the surface. Contours of positive maxima of κ_1 and contours or negative minima of κ_1 swap places and the partitioning along the new negative minima becomes apparent.

As noted in the introduction, segmentation rules which use only the Gaussian curvature, rather than analyzing the principal curvatures independently, fail on this example and on cones because the Gaussian curvature is everywhere zero, making impossible any segmentation based only upon the Gaussian curvature. Yet human observers readily and consistently perceive partitions in surfaces whose Gaussian curvature is everywhere zero.

Cones

Cones are a special case of ruled surfaces in which the directrix, $\mathbf{p}$, is simply a point, the vertex of the cone. In consequence one can give the following parametrization for the cone:

$$\mathbf{x}(u^1, u^2) = u^2\mathbf{v}(u^1), \quad u^1 \in (a, b) \subset \mathbf{R}, \quad u^2 \in \mathbf{R}.$$

For this parametrization
$\mathbf{x}_2 = \mathbf{v}$. The metric tenso

g_{ij} =

The surface normal is

$\mathbf{N} = \frac{\mathbf{x}_1}{|\mathbf{x}_1}$

The second fundamenta

$b_{ij} = \mathbf{x}_{ij}$

Since $g_{12} = b_{12} = 0$

$\kappa_1 = b_{11}/g_{11}$

The u^1- and u^2-parame
are both zero. As one w
u^2-parameter curves is ev
parameter curves (where
tours of negative minima
of the cone. An example c
tours indicated by dotted l

Figure 14. Partitions of a

$+ \begin{pmatrix} |\mathbf{p}_{11}| & 0 \\ 0 & 0 \end{pmatrix}$

on a cylinder are

second derivative of **p** with
he orientation of the field of
along the directrix **p**. The
of lines of curvature and the
. As expected, the curvature
o partitioning contours arise
the principal curvature κ_2).
its translations are used for

ows a cylinder and its parti-
ientations of the field of sur-
cylinder into parts that seem
ll experience a figure-ground
become dips and vice-versa.
se the perceived partitioning
nd to the lines that were pre-
ise the figure-ground reversal
the field of surface normals
surface. Contours of positive
f κ_1 swap places and the par-
apparent.
les which use only the Gaus-
pal curvatures independently,
Gaussian curvature is every-
n based only upon the Gaus-
d consistently perceive parti-
rywhere zero.

ich the directrix, **p**, is simply
one can give the following

$) \subset \mathbf{R}, \quad u^2 \in \mathbf{R}.$

For this parametrization the first partial derivatives are $\mathbf{x}_1 = u^2\mathbf{v}_1$ and $\mathbf{x}_2 = \mathbf{v}$. The metric tensor is

$$g_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j = \begin{pmatrix} (u^2)^2(\mathbf{v}\cdot\mathbf{v}) & 0 \\ 0 & 1 \end{pmatrix}$$

The surface normal is

$$\mathbf{N} = \frac{\mathbf{x}_1 \times \mathbf{x}_2}{|\mathbf{x}_1 \times \mathbf{x}_2|} = \frac{u^2(\mathbf{v}_1 \times \mathbf{v})}{u^2|\mathbf{v}_1 \times \mathbf{v}|} = \frac{\mathbf{v}_1 \times \mathbf{v}}{|\mathbf{v}_1|}$$

The second fundamental coefficients are

$$b_{ij} = \mathbf{x}_{ij} \cdot \mathbf{N} = \begin{pmatrix} u^2(\mathbf{v}_1, \mathbf{v}, \mathbf{v}_{11})/|\mathbf{v}_1| & 0 \\ 0 & 0 \end{pmatrix}$$

Since $g_{12} = b_{12} = 0$ the principal curvatures on a cone are

$$\kappa_1 = b_{11}/g_{11} = \frac{u^2(\mathbf{v}_1, \mathbf{v}, \mathbf{v}_{11})}{(u^2)^2(\mathbf{v}_1\cdot\mathbf{v}_1)|\mathbf{v}_1|} = \frac{(\mathbf{v}_1, \mathbf{v}, \mathbf{v}_{11})}{u^2|\mathbf{v}_1|^3}$$

$$\kappa_2 = b_{22}/g_{22} = 0$$

The u^1- and u^2-parameter curves are lines of curvature, since g_{12} and b_{12} are both zero. As one would expect, the principal curvature, κ_2, along the u^2-parameter curves is everywhere zero. The expression for κ_1 along the u^1-parameter curves (where u^2 is constant) does not depend on u^2. Thus the contours of negative minima of κ_1 are straight lines which radiate from the vertex of the cone. An example cone is shown in Figure 14 with the partitioning contours indicated by dotted lines. The resulting parts appear quite natural.

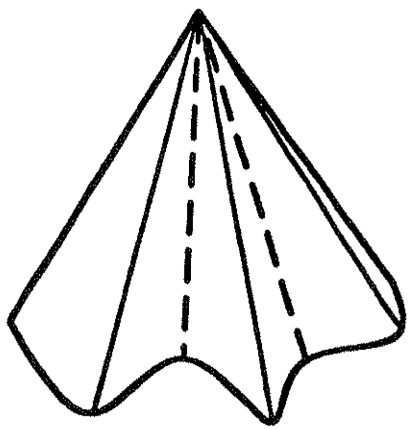


Figure 14. Partitions of a cone

Surfaces of revolution

A surface of revolution is a set $S \subset \mathbf{R}^3$ obtained by rotating a regular plane curve, α, about an axis in the plane which does not meet the curve. Let the xz-plane be the plane of α and let the axis of rotation be the z-axis. Let

$$\alpha(u^1) = (x(u^1), z(u^1)), \quad a < u^1 < b, \quad x(u^1) > 0,$$

and let u^2 be the rotation angle about the z-axis. Then we obtain a map

$$\mathbf{x}(u^1, u^2) = (x(u^1)\cos(u^2), x(u^1)\sin(u^2), z(u^1))$$

from the open set $U = \{(u^1, u^2) \in \mathbf{R}^2;\ 0 < u^2 < 2\pi,\ a < u^1 < b\}$ into S (as shown in Figure 15). The curve α is called the *generating curve* of S, and the z-axis is the *rotation axis* of S. The circles swept out by the points of α are called the *parallels* of S, and the various placements of α on S are called the *meridians* of S.

The partial derivatives are $\mathbf{x}_1 = (x_1 \cos(u^2),\ x_1 \sin(u^2),\ z_1)$ and $\mathbf{x}_2 = (-x \sin(u^2), x\cos(u^2), 0)$. The metric tensor is

$$g_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j = \begin{pmatrix} x_1^2 + z_1^2 & 0 \\ 0 & x^2 \end{pmatrix}.$$

The surface normal is

$$\mathbf{N} = \frac{\mathbf{x}_1 \times \mathbf{x}_2}{|x_1 \times x_2|} = \frac{(z_1 \cos(u^2),\ z_1 \sin(u^2),\ -x_1)}{\sqrt{z_1^2 + x_1^2}}.$$

If we let u^1 be arc length along α then $\sqrt{z_1^2 + x_1^2} = 1 = g_{11}$ and

$$\mathbf{N} = (z_1 \cos(u^2),\ z_1 \sin(u^2),\ -x_1).$$

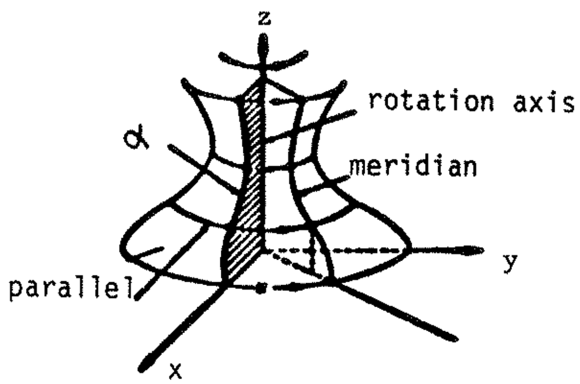


Figure 15. A convenient parametrization for a surface of revolution

The second fundamental coe

$b_{ij} = \mathbf{x}_{ij}\cdot$

Since $g_{12} = b_{12} = 0$ the p

$\kappa_1 =$

κ

The expression for κ_1 is
α. In fact the meridians (th
ture, as are the parallels. T
expression for κ_1 and the cu
sion for κ_2. The expression
x multiplied by the cosine
axis of rotation.

Observe that the express
u^1, not u^2. In particular, si
inflections of the normal cu
cles. Consequently no parti
associated with κ_2. Only th
partitioning. Figure 16 sho
of curvature along the meri
appear natural.

Figure 1, as discussed
orientation (of the field of
us to carve the same surfac

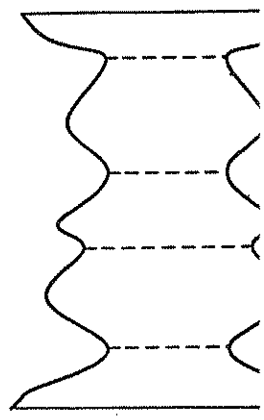


Figure 16. Partitions of som

by rotating a regular plane
ıot meet the curve. Let the
on be the z-axis. Let

$b, \quad x(u^1) > 0,$

hen we obtain a map

$\text{in}(u^2), \ z(u^1))$

$< 2\pi, \ a < u^1 < b\}$ into S
generating curve of S, and
vept out by the points of α
ements of α on S are called

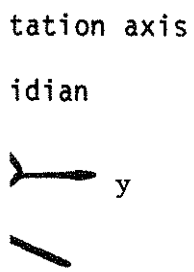
$u^2), \ x_1 \sin(u^2), \ z_1)$ and
or is

$\begin{bmatrix} 0 \\ x^2 \end{bmatrix}.$

$\frac{\sin(u^2), \ -x_1)}{+ \ x_1^2}.$

$= 1 = g_{11}$ and

$-x_1).$



of revolution

The second fundamental coefficients are

$$b_{ij} = \mathbf{x}_{ij} \cdot \mathbf{N} = \begin{bmatrix} x_{11} z_1 - s_1 z_{11} & 0 \\ 0 & -x z_1 \end{bmatrix}.$$

Since $g_{12} = b_{12} = 0$ the principal curvatures on a surface of revolution are

$$\kappa_1 = b_{11}/g_{11} = x_{11} z_1 - x_1 z_{11},$$

$$\kappa_2 = b_{22}/g_{22} = -z_1/x.$$

The expression for κ_1 is identical to the expression for the curvature along α. In fact the meridians (the various positions of α on S) are lines of curvature, as are the parallels. The curvature along the meridians is given by the expression for κ_1 and the curvature along the parallels is given by the expression for κ_2. The expression for κ_2 is simply the curvature of a circle of radius x multiplied by the cosine of the angle that the tangent to α makes with the axis of rotation.

Observe that the expressions for κ_1 and κ_2 depend only upon the parameter u^1, not u^2. In particular, since κ_2 is independent of u^2 there are no extrema or inflections of the normal curvature along the parallels. The parallels are circles. Consequently no partitioning contours arise from the lines of curvature associated with κ_2. Only the minima of κ_1 along the meridians are used for partitioning. Figure 16 shows several surfaces of revolution with the minima of curvature along the meridians marked. The resulting partitioning contours appear natural.

Figure 1, as discussed in the introduction, illustrates that reversing the orientation (of the field of surface normals) of a surface of revolution causes us to carve the same surface differently. The dotted circular lines in the figure

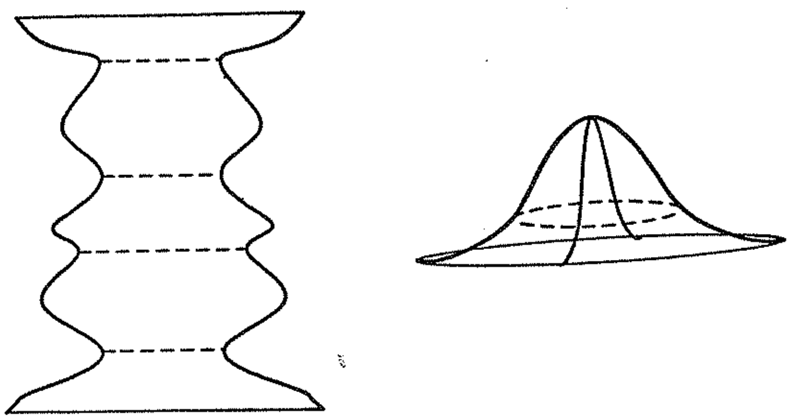


Figure 16. Partitions of some surface of revolution

are the partitioning contours according to the negative minima rule. Note that they lie in the valleys of the top figure. If the figure is inverted they no longer lie in the valleys but on the peaks. By reversing the field of surface normals the signs of the principal curvatures everywhere have reversed. Contours of negative minima of the principal curvatures become contours of positive maxima, and vice-versa. Consequently the part boundaries are not invariant under a reversal of orientation.

The torus

A torus is a surface in $\mathbf{R}^3$ which is obtained by revolving a circle about a line not passing through the circle, as shown in Figure 17. A convenient parametrization for the torus is

$$\mathbf{x}(u^1, u^2) = ((b + a\sin(u^2))\cos(u^1), (b + a\sin(u^2))\sin(u^1), a\cos(u^2)),$$

$$b > a.$$

The first partials are $x_1 = (-(b + a\sin(u^2))\cos(u^1),\ (b + a\sin(u^2))\cos(u^1),\ 0)$ and $x_2 = (a\cos(u^2)\cos(u^1),\ a\cos(u^2)\sin(u^1),\ -a\sin(u^2))$. The metric tensor is

$$g_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j = \begin{pmatrix} (b + a\sin(u^2))^2 & 0 \\ 0 & a^2 \end{pmatrix}.$$

The surface normal is

$$\mathbf{N} = \frac{\mathbf{x}_1 \times \mathbf{x}_2}{|\mathbf{x}_1 \times \mathbf{x}_2|} = (-\cos(u^1)\sin(u^2),\ -\sin(u^1)\sin(u^2),\ -\cos(u^2)).$$

The second fundamental coefficients are

$$b_{ij} = \mathbf{x}_{ij} \cdot \mathbf{N} = \begin{pmatrix} (b + a\sin(u^2))\sin(u^2) & 0 \\ 0 & a \end{pmatrix}.$$

Since $g_{12} = b_{12} = 0$ the u^1- and u^2-parameter curves are lines of curvature and the principal curvatures on a torus are

$$\kappa_1 = b_{11}/g_{11} = \sin(u^2)/(b + a\sin(u^2)),$$

$$\kappa_2 = b_{22}/g_{22} = a^{-1}.$$

The principal curvature κ_1 is associated with the u^1-parameter curves and κ_2 with the u^2-parameter curves. κ_2 is a constant, so the torus is not parti-

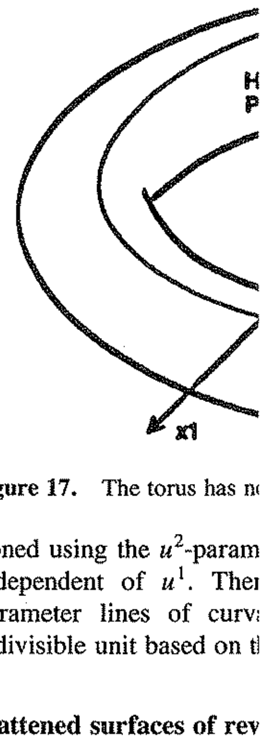


Figure 17. The torus has n

tioned using the u^2-param
independent of u^1. Ther
parameter lines of curv
indivisible unit based on t

Flattened surfaces of rev

What happens to the par
flatten it slightly along on
here that the circular parti
elliptical and bowed sligh
to test this against percept
Figure 18 illustrates a
tion which is flattened:

$\mathbf{x}(u^1, u^2) = (f(u^1)$

Let $f(u^1)$ be abbreviat
tive with respect to u^1. Th

$g_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j$

$= \Big((f')^2(\cos^2(u^2)$ ... $ff'\sin(u^2)\mathrm{cc}$

: negative minima rule. Note that
: figure is inverted they no longer
:sing the field of surface normals
'here have reversed. Contours of
become contours of positive max-
oundaries are not invariant under

by revolving a circle about a line
'igure 17. A convenient parametr-

$a \sin(u^2)) \sin(u^1), a \cos(u^2)),$

$in(u^2)) \cos(u^1), \quad (b + a \sin(u^2))$
$a \cos(u^2) \sin(u^1), \quad -a \sin(u^2)).$

$in(u^2))^2 \quad 0$
$) \quad a^2$

$- \sin(u^1) \sin(u^2), \; - \cos(u^2)).$

$(u^2)) \sin(u^2) \quad 0$
$0 \quad a$

ameter curves are lines of curva-
ıre

$b + a \sin(u^2)),$

$a^{-1}.$

with the u^1-parameter curves and
onstant, so the torus is not parti-

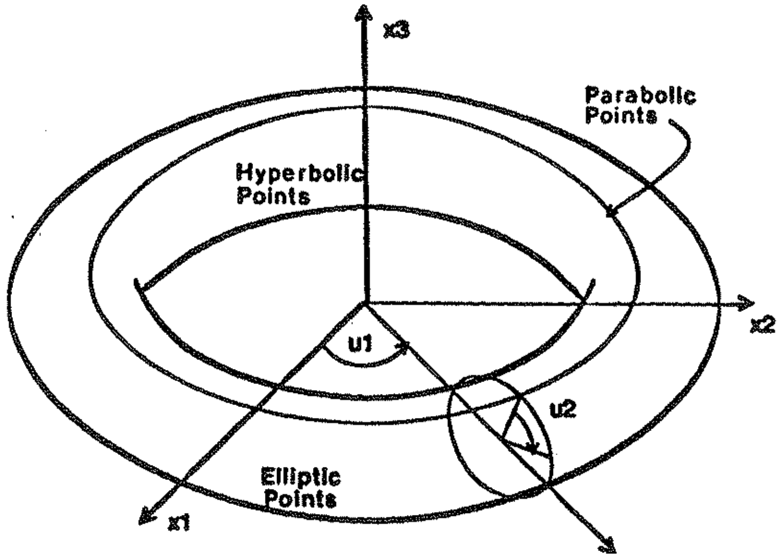


Figure 17. The torus has no parts

tioned using the u^2-parameter lines of curvature. κ_1 is not a constant, but it is independent of u^1. Therefore the torus is not partitioned using the u^1-parameter lines of curvature either. We conclude that the torus is one indivisible unit based on the negative minima partitioning rule.

Flattened surfaces of revolution

What happens to the partitioning contours on a surface of revolution if we flatten it slightly along one axis orthogonal to the axis of revolution? We show here that the circular partitioning contours of the surface of revolution become elliptical and bowed slightly up or down in the middle. It would be of interest to test this against perceptual judgments.

Figure 18 illustrates a convenient parametrization for a surface of revolution which is flattened:

$$\mathbf{x}(u^1, u^2) = (f(u^1)\cos(u^2),\ af(u^1)\sin(u^2),\ (u^1)), \quad 0 < a < 1.$$

Let $f(u^1)$ be abbreviated to f and let primes over the f's indicate derivative with respect to u^1. Then the metric tensor is

$$g_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j$$

$$= \begin{bmatrix} (f')^2(\cos^2(u^2) + a^2 \sin^2(u^2)) + 1 & f f' \sin(u^2)\cos(u^2)(a^2 - 1) \\ f f' \sin(u^2)\cos(u^2)(a^2 - 1) & f^2(\sin^2(u^2) + a^2\cos^2(u^2)) \end{bmatrix}$$

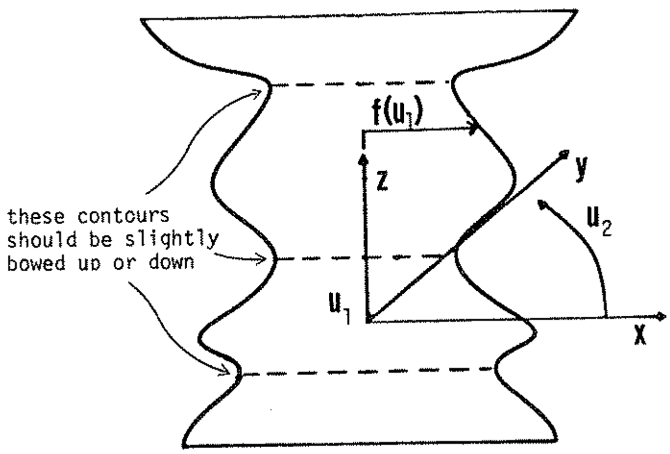


Figure 18. Partitions on a flattened surface of revolution

The second fundamental coefficients are

$$b_{ij} = \mathbf{x}_{ij}\cdot\mathbf{N} = \begin{pmatrix} -af''/d & 0 \\ 0 & af/d \end{pmatrix}$$

where $d = \sqrt{a^2\cos^2(u^2) + \sin^2(u^2)} = a^2(f')^2$.

Since $\mathbf{x}_1\cdot\mathbf{x}_2 \neq 0$ in general, the parameter curves are not in general lines of curvature. However when $f' = 0$ then $\mathbf{x}_1\cdot\mathbf{x}_2 = 0$ so that contours where this holds are lines of curvature. These contours are elliptical cross sections of the flattened surface of revolution, cross sections having either the greatest or least major axis locally. Along these lines of curvature the associated principal curvature is

$$\kappa = b_{22}/g_{22} = af^{-1}(\sin^2(u^2) + a^2\cos(u^2))^{-3/2}.$$

Its extrema occur when

$$\partial\kappa/\partial u^2 = -7/2af^{-1}(a^2\cos^2(u^2) + \sin^2(u^2))^{-5/2}(2\cos(u^2)\sin(u^2) - 2a^2\cos(u^2)\sin(u^2)) = 0,$$

which happens when $a^2\cos(u^2)\sin(u^2) = \cos(u^2)\sin(u^2)$. This implies that $u^2 = n\pi/2$, for n an integer. For n even, $\kappa = a^{-2}f^{-1}$, and for n odd, $\kappa = af^{-1}$. Thus the minima occur when u^2 is $\pi/2$ or $3\pi/2$. However these minima are positive minima, since a and f are both positive, and consequently there are no partitioning contours which arise from this family of lines of curvature.

To determine the pa of curvature, we begir becomes

for n even, and

for n odd. Thus the u^1-p vature. These curves ar lution with the xz-plane vature is

$\kappa =$

and for n odd it is

$\kappa =$

The extrema of these values of u^1 (because a surface of revolution are is flattened, the partitio elliptical and usually bo

Elbows

An apparent problem "elbow", the problem b of curvature are not clo for smooth elbows and,

As can be seen in Fig specify an incomplete p the partition is inherentl in the figure is equally re

Elongated torus

Elbows may also occur which has been scaled a

To determine the partitioning contours defined by the other family of lines of curvature, we begin by noting that when u^2 is $n\pi/2$ the metric tensor becomes

$$g_{ij} = \begin{pmatrix} (f')^2 + 1 & 0 \\ 0 & a^2 f^2 \end{pmatrix}$$

for n even, and

$$g_{ij} = \begin{pmatrix} a^2 (f')^2 + 1 & 0 \\ 0 & f^2 \end{pmatrix}$$

for n odd. Thus the u^1-parameter curves given by $u^2 = n\pi/2$ are lines of curvature. These curves are also the intersection of the flattened surface of revolution with the xz-plane or yz-plane. For n even the associated principal curvature is

$$\kappa = b_{11}/g_{11} = -f''(1 + (f')^2)^{-3/2},$$

and for n odd it is

$$\kappa = b_{11}/g_{11} = -af''(1 + a(f')^2)^{-3/2}.$$

The extrema of these two curvatures do not, in general, occur at the same values of u^1 (because $a \neq 1$). Thus the partitioning contours on the flattened surface of revolution are not, in general, planar. So as a surface of revolution is flattened, the partitioning contours which are at first circles become more elliptical and usually bow either up or down slightly.

Elbows

An apparent problem for the negative minima partitioning rule is the "elbow", the problem being that for elbows the contours of negative minima of curvature are not closed, so no parts are uniquely delimited. This is true for smooth elbows and, as shown in Figure 19, for nonsmooth elbows as well.

As can be seen in Figure 19, however, there is good reason for the rule to specify an incomplete partitioning contour—the appropriate way to continue the partition is inherently ambiguous. Each of these three completions shown in the figure is equally reasonable.

Elongated torus

Elbows may also occur on entirely smooth surfaces. For instance, the torus which has been scaled along one axis has two elbows. The following deriva-

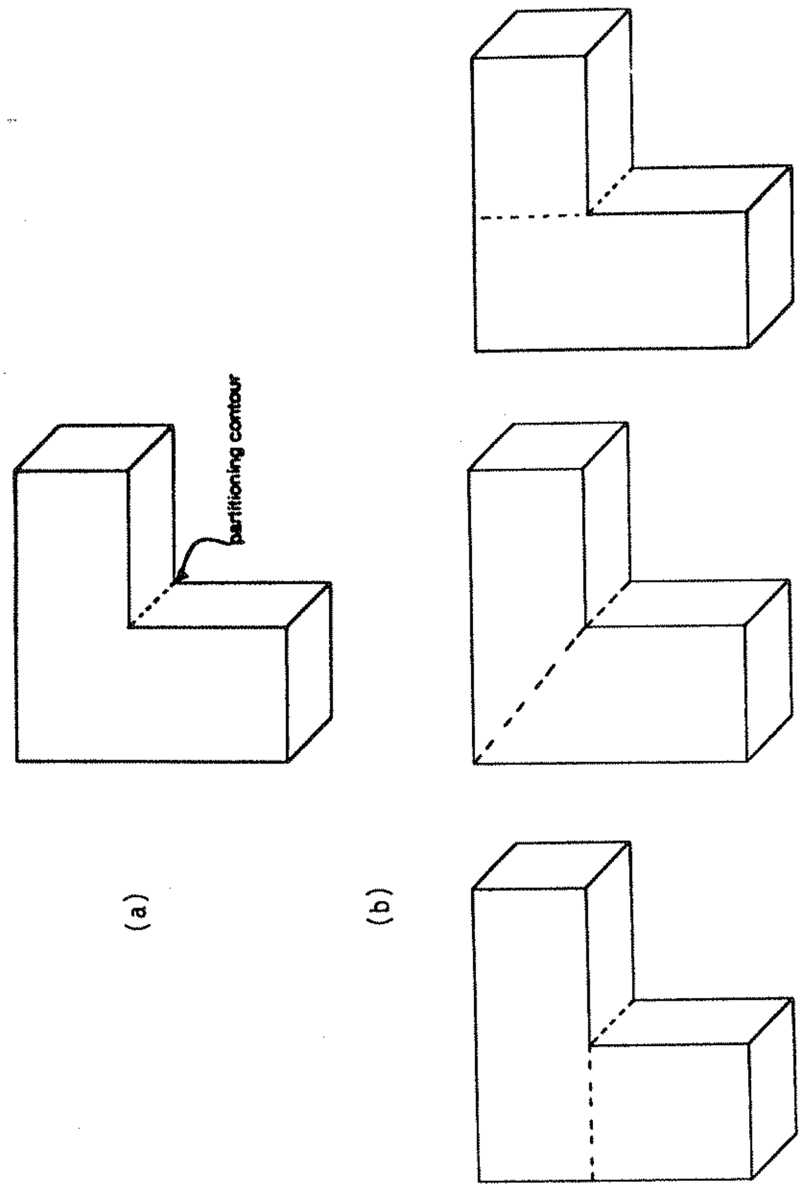


Figure 19. Partitioning of an elbow

tion will show that the negat
lar contours, one on the insi

The elongated torus may

$$\mathbf{x}(u^1, u^2)$$

$$= ((b + a \sin(u^2)) \cos(u$$

This corresponds in Figure
first partials are

$$\mathbf{x}_1 = (-(b + a\sin($$

and

$$\mathbf{x}_2 = (a \cos(u^2)\ c$$

The metric tensor is

$$\left[\begin{array}{l} (b + a \sin(u^2))^2(\sin^2(u^1 + \\ a \cos(u^2) \cos(u^1) \sin(u^1)(b + a \end{array} \right.$$

The surface normal is

$$\mathbf{N} = (-d \sin(u^2)\ co$$

where

$$f = \sqrt{d^2 \sin^2(u^2)\ cos$$

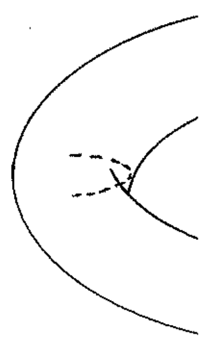


Figure 20. An elongated torus

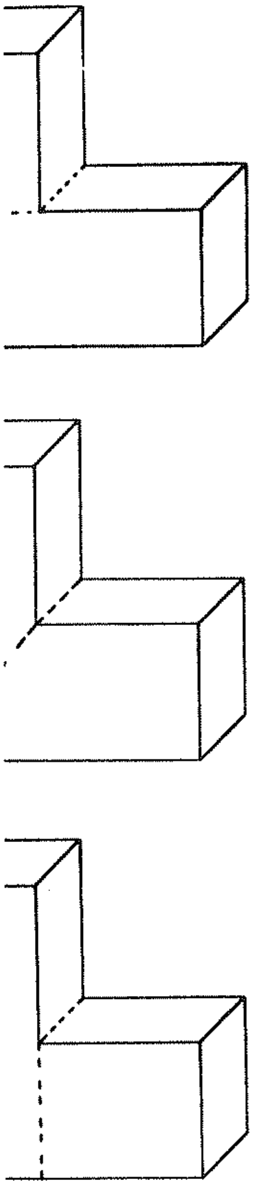


Figure 19. Partitioning of an elbow

tion will show that the negative minima rule gives rise to two open semicircular contours, one on the inside of each elbow, as shown in Figure 20.

The elongated torus may be conveniently parametrized as

$$\mathbf{x}(u^1, u^2)$$

$$= ((b + a\sin(u^2))\cos(u^1),\ d(b + a\sin(u^2))\sin(u^1),\ a\cos(u^2)), \quad b > a,\ d > 1.$$

This corresponds in Figure 17 to expanding the torus along the x^2-axis. The first partials are

$$\mathbf{x}_1 = (-(b + a\sin(u^2))\sin(u^1), d(b + a\sin(u^2))\cos(u^1), 0)$$

and

$$\mathbf{x}_2 = (a\cos(u^2)\cos(u^1), ad\cos(u^2)\sin(u^1), -a\sin(u^2)).$$

The metric tensor is

$$\begin{pmatrix} (b + a\sin(u^2))^2(\sin^2(u^1 + d^2\cos^2(u^1)) & a\cos(u^2)\cos(u^1)\sin(u^1)(b + a\sin(u^2))(d^2 - 1) \\ a\cos(u^2)\cos(u^1)\sin(u^1)(b + a\sin(u^2))(d^2 - 1) & a^2(\cos^2(u^2)\cos^2(u^1) + d^2\cos^2(u^2)\sin^2(u^1) + \sin^2(u^2)) \end{pmatrix}$$

The surface normal is

$$\mathbf{N} = (-d\sin(u^2)\cos(u^1),\ -\sin(u^1)\sin(u^2),\ -d\cos(u^2))/f,$$

where

$$f = \sqrt{d^2\sin^2(u^2)\cos^2(u^1) + \sin^2(u^1)\sin^2(u^2) + d^2\cos^2(u^2)}.$$

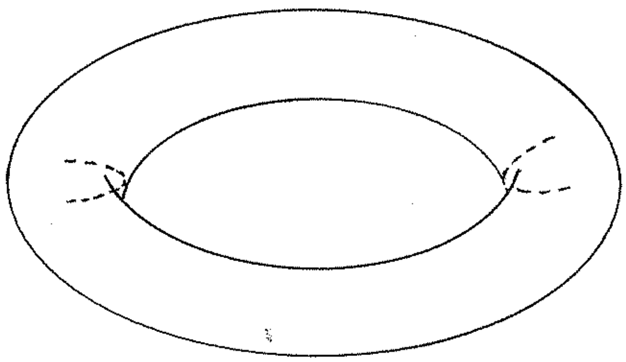


Figure 20. An elongated torus has two semi-circular contours of partition

The second fundamental coefficients are

$$b_{ij} = \begin{bmatrix} d\sin(u^2)(b + a\sin(u^2))/f & 0 \\ 0 & ad/f \end{bmatrix}.$$

Since $g_{12} \neq 0$ the u^1- and u^2-parameter curves are not in general lines of curvature. However, along the curve $u^2 = \pi/2$ we have $\cos(u^2) = 0$, $\sin(u^1) = 1$, and

$$g_{ij} = \begin{bmatrix} (b + a)^2(\sin^2(u^1) + d^2\cos^2(u^1)) & 0 \\ 0 & a^2 \end{bmatrix},$$

implying that this is a line of curvature (b_{12} and g_{11} are zero). The second fundamental coefficients are

$$b_{ij} = \begin{bmatrix} d(b + a)/h & 0 \\ 0 & ad/h \end{bmatrix},$$

where

$$h = \sqrt{d^2\cos^2(u^1) + \sin^2(u^1)}.$$

The principal curvature along this line of curvature is

$$\kappa = b_{11}/g_{11} = d(b + a)^{-1}h^{-3}.$$

The extrema of curvature along this line occur where $\partial\kappa/\partial u^1 = 0$.

$$\frac{\partial\kappa}{\partial u^1} = -3/2d(b+a)^{-1}(d^2\cos^2(u^1) + \sin^2(u^1))^{-5/2}(2\sin(u^1)\cos(u^1) - 2d^2\sin(u^1)\cos(u^1)) = 0.$$

Since $d > 1$ this implies that $d^2\cos(u^1)\sin(u^1) = \cos(u^1)\sin(u^1)$, which occurs for $u^1 = n\pi/2$, where n is an integer. Positive maxima of curvature occur when n is odd, negative minima when n is even.

A similar analysis shows that the contour $u^2 = -\pi/2$ is a line of curvature whose extrema occur for $u^1 = n\pi/2$, where n is an integer. The difference is that the positive maxima of curvature occur when n is even, negative minima when n is odd.

Finally, at the parameter point $(\pi/2, 0)$ we have that $\sin(u^1 = 1$, $\cos(u^1) = 0$, $\sin(u^2) = 0$, $\cos(u^2) = 1$, and find that the metric tensor is

implying that at this poin
directions. The second fun

Hence $\kappa_1 = b_{11}/g_{11} = 0$.
ter points $(-\pi/2, 0)$, (
then, we have found tha
minimum of κ_1, the oute
lowermost points have κ_1
tioning contours at the
$\pi < u^2 < 0$, and $u^1 =$

5. SUMMARY

To recognize an object fr
into parts. Defining part
afford the broadest possib
generic, stable property
boundary-based partitioni
that when one smooths a
trarily large curvature as
solid union, positive for s
propose, in consequence,
pal curvatures and some
tures are used by the h
which tell when to use
minima are being devel
theory to multiple scales

REFERENCES

Barr, A. (1983). Ph.D. Dissert
Blinn, J. (1982). A generalizat
1(3), 235–256.
Brady, J.M., & Asada, H. (1
Memo 757). Cambridge

$$g_{ij} = \begin{bmatrix} b^2 & 0 \\ 0 & a^2d^2 \end{bmatrix},$$

implying that at this point the u^1- and u^2-parameter curves are in principal directions. The second fundamental form is

$$b_{ij} = \begin{bmatrix} 0 & 0 \\ 0 & ad \end{bmatrix}.$$

Hence $\kappa_1 = b_{11}/g_{11} = 0$. By symmetry this also holds for κ_1 at the parameter points $(-\pi/2, 0)$, $(-\pi/2, \pi)$, $(\pi/2, \pi)$. At each of the two elbows, then, we have found that the innermost point of the elbow is a negative minimum of κ_1, the outermost is a positive maximum, the uppermost and lowermost points have $\kappa_1 = 0$. By symmetry we conclude that the two partitioning contours at the elbows are the open semicircles $u^1 = \pi/2$, $\pi < u^2 < 0$, and $u^1 = -\pi/2$, $\pi < u^2 < 0$.

5. SUMMARY

To recognize an object from its shape it is useful first to decompose the shape into parts. Defining parts by their boundaries, rather than by their shapes, afford the broadest possible scope to the partitioning scheme. Transversality, a generic, stable property of the intersection of surfaces, motivates the boundary-based partitioning scheme considered here. In particular we show that when one smooths a transversal intersection of surfaces one obtains arbitrarily large curvature as the intersection curve is approached (negative for solid union, positive for solid subtraction) regardless of how one smooths. We propose, in consequence, that some contours of negative minima of the principal curvatures and some contours of positive maxima of the principal curvatures are used by the human visual system as part boundaries. The rules which tell when to use positive maxima and when instead to use negative minima are being developed. Also to be developed is an extension of this theory to multiple scales of resolution.

REFERENCES

Barr, A. (1983). Ph.D. Dissertation, Rensselaer Polytechnic Institute, Troy, NY.

Blinn, J. (1982). A generalization of algebraic surface drawing. *ACM Transactions on Graphics, 1*(3), 235–256.

Brady, J.M., & Asada, H. (1984). *Smoothed local symmetries and their implications* (MIT AI Memo 757). Cambridge, MA: MIT.

Do Carmo, M. (1976). *Differential geometry of curves and surfaces*, Englewood Cliffs, NJ: Prentice-Hall.
Hoffman, D. (1983). The interpretation of visual illusions. *Scientific American, 249*(6), 154–162.
Hoffman, D., & Richards, W. (1984). Parts of recognition. *Cognition, 18*, 65–96.
Koenderink, J., & van Doorn, A. (1980). Photometric invariants related to solid shape. *Optica Acta, 7*, 981–996.
Koenderink, J., & van Doorn, A. (1982). The shape of objects and the way contours end. *Perception, 11*, 129–137.

chapter 9

An interna representa shape bas topologica the appare

1. INTRODUCTION

The problem of recognizin
of "internal representation
eye's view can be compa
sional, in the sense that a
possible to construct a set
the object with sufficient a
about such generic 2D qua

An example may illustr
see at a single glance whe
way all views of the fro
roughly only through a pr
when you take a few step
back views differ qualitat
separately. I call the quali
the "parts" of an object i

* This work was supported i