a2 Geometry: The Local Theory of Curves and Surfaces.

Richard Earl

Trinity 1998

Contents

0.0 Introduction

The aim of the a2 geometry course is to introduce the student to curves and surfaces and to some elementary local aspects of theory.

Simplistically a curve or a surface is a space that 'looks' locally like a line or a plane. More generally geometers consider higher dimensional 'manifolds' which locally resemble higher dimensional Euclidean space. However high the dimension the fundamentals of local geometric theory remain the same.

Saying that a space is locally Euclidean means that around each point we can introduce some co-ordinates – one co-ordinate, such as arc-length, for a curve; two co-ordinates, like the x and y co-ordinates in the plane, for a surface, and so on. With a notion of co-ordinates local geometry on a curve or surface is then no more complicated than the usual study of Euclidean space and boils down to the calculus of one or two variables.

All the above is classical differential geometry and appears in most of the standard texts. For example:

Differential Geometry of Curves and Surfaces
M.P. do Carmo, Prentice Hall (1976).

Geometry of Surfaces
G. Segal, Mathematical Institute Notes (1986).

Whilst the above two texts are both excellent their aims are far grander and deeper than the scope and syllabus of this brief introductory course and this is very much the motivation for writing these notes. These notes discuss primarily the theory in the a2 geometry syllabus. They are intended as a leisurely introduction to this theory and contain more examples and exercises than most students may need. The notes also contain aspects of geometric theory which are not explicitly mentioned in the syllabus. This is partly done to put the geometry into better context and also to suggest further reading where appropriate. At these stages the relevant theorem, definition etc. is marked by an asterisk.

I hope these notes prove useful to understanding the a2 course. For those interested in geometry the b3 course (Geometry of Surfaces) is a natural sequel. I would also welcome any comments on these notes and especially any suggestions on how they may be improved or amended.

Richard Earl,
Mathematical Institute,
earl@maths.ox.ac.uk,
September 1998.

Chapter 1

Curves

1.1 Parametrised Curves

DEFINITION 1 *A* smooth parametrised curve *in* $\mathbf{R}^3$ *is a map* $\gamma : I \to \mathbf{R}^3$ *from an open interval* $I \subseteq \mathbf{R}$ *such that*

- γ *is smooth i.e.* $\gamma(t) = (x(t), y(y), z(t))$ *where* x, y, z *have derivatives of all orders,*
- $\gamma : I \to \gamma(I)$ *is a homeomorphism,*
- $\gamma'(t) \neq \mathbf{0}$ *for all* $t \in I$.

The requirement that γ be a homeomorphism onto its image is somewhat unusual here. Some authors will omit this requirement which allows the possibility of self-intersections, for the curve crossing itself. Defining a smooth parametrised curve as above means that the curve has no singular points and also mirrors the definition of a smooth parametrised surface (see §2.1).

A smooth parametrised curve γ is a curve in $\mathbf{R}^3$ with a preferred parametrisation. The image of γ is also the image of other smooth parametrised curves. Throughout these notes we will need to check that definitions we make for curves and surfaces are independent of the choice of parametrisation. For example, a simple application of the chain rule shows that the tangent line to a curve and arc-length on a curve (as defined below) are independent of the choice of parameter. Arc-length is a 'natural' parameter for a curve.

DEFINITION 2 *Let* $\gamma : I \to \mathbf{R}^3$ *be a smooth parametrised curve and let* $t_0 \in I$. *Then the* arc-length $s(t)$ *from* $\gamma(t_0)$ *to a point* $\gamma(t)$ *is defined to be the integral,*

$$s(t) = \int_{t_0}^{t} |\gamma'(u)| \, \mathrm{d}u.$$

As $\gamma'(t) \neq \mathbf{0}$ for all t then there is a well defined *tangent line* at each point of $\gamma(I)$.

DEFINITION 3 *Let* $\gamma : I \to \mathbf{R}^3$ *be a smooth parametrised curve and let* $t_0 \in I$. *Then the* tangent line *to* γ *at* $\gamma(t_0)$ *is the line containing the point* $\gamma(t_0)$ *in the direction* $\gamma'(t_0)$. *The* unit tangent vector *is the tangent vector* $\mathbf{t} = \mathrm{d}\gamma/\mathrm{d}s$.

Example 4 *Q: Let* $\gamma(t) = (ae^{bt}\cos t, ae^{bt}\sin t)$ *for* $t \in \mathbf{R}$ *and real constants* $a > 0 > b$. *Show that* γ *has finite arc-length in* (t_0, ∞) *for any* $t_0 \in \mathbf{R}$.

A: The tangent vector $\gamma'(t)$ equals

$$(ae^{bt}(b\cos t - \sin t), ae^{bt}(b\sin t + \cos t)),$$

and has magnitude

$$ae^{bt}\sqrt{((b\cos t - \sin t)^2 + (b\sin t + \cos t)^2)} = ae^{bt}\sqrt{(b^2+1)}.$$

So the arc-length from $\gamma(t_0)$ to $\lim_{t\to\infty}\gamma(t) = (0,0)$ equals

$$a\sqrt{1+b^2}\int_{t_0}^{\infty} e^{bu}\mathrm{d}u = a\sqrt{1+b^{-2}}e^{bt_0}. \qquad \square$$

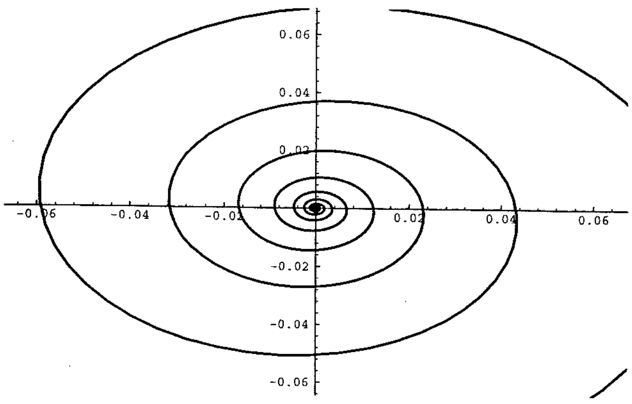


Figure 1 – The Logarithmic Spiral ($a = 1, b = -0.1$)

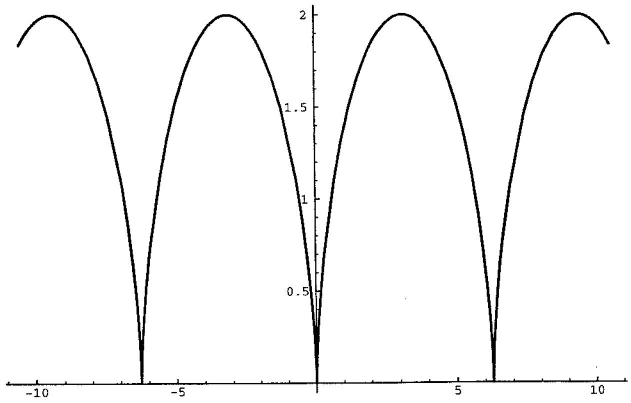


Figure 2 – The Cycloid ($r = 1$)

Example 5 *Q: The* tractrix *is the curve given by*

$$\gamma(t) = (-(\cos t + \log \tan \frac{t}{2}), \sin t) \qquad t \in (0, \pi/2).$$

Show that the length of the tangent line from a point $\gamma(t)$ to where the tangent meets the x-axis is always 1.

A: Differentiating we find that $\gamma'(t)$ equals

$$\left(\frac{-\cos^2 t}{\sin t}, \cos t\right), \qquad t \in (0, \pi/2).$$

So the tangent from the curve at $\gamma(t)$ meets the x-axis at

$$\gamma(t) + (\cos t, -\sin t),$$

a point distance 1 away. $\square$

Example 6 *Q: A circular disc of radius r in the xy-plane rolls without slipping along the x-axis. The figure described by a point of the circumference of the disc is called a* cycloid *(see Figure 2). Determine the arc-length of a section of the cycloid which corresponds to a complete rotation of the disc.*

A: Assume that the disc begins with its centre at $(0, r)$. Consider the curve described by the point $(0, 0)$ as the disc rolls.

After the disc has rolled distance $r\theta$ then the point $(0, 0)$ has moved on to

$$(x(\theta), y(\theta)) = (r(\theta - \sin\theta), r(1 - \cos\theta)).$$

Thus $x'(\theta)^2 + y'(\theta)^2 = 2r^2(1 - \cos\theta)$ and so

$$s = \sqrt{2}\, r \int_0^{2\pi} \sqrt{1 - \cos\theta}\, \mathrm{d}\theta = 2r \int_0^{2\pi} \left|\sin \frac{1}{2}\theta\right| \mathrm{d}\theta = 8r. \quad \square$$

1.2 The Serret-Frenet Formulae

DEFINITION 7 *Let $\gamma : I \to \mathbf{R}^3$ be a smooth curve parametrised by arc-length s. Let*

$$\mathbf{t}(s) = \frac{\mathrm{d}\gamma}{\mathrm{d}s}$$

denote the unit tangent to γ at s. For a smooth function x of s we shall write $\dot{x}$ for $\mathrm{d}x/\mathrm{d}s$.

(a) The curvature *$\kappa(s)$ of γ at s is defined to be*

$$\kappa(s) = |\ddot{\gamma}(s)| = |\dot{\mathbf{t}}(s)|.$$

(b) If $\kappa(s) \neq 0$ then the normal $\mathbf{n}(s)$ *to γ at s is the unit vector determined by the equation*

$$\dot{\mathbf{t}}(s) = \kappa(s)\mathbf{n}(s).$$

If $\kappa(s) = 0$ then $\mathbf{n}(s)$ is not defined.

(c) The plane in $\mathbf{R}^3$ containing $\gamma(s)$ and the vectors $\mathbf{t}(s)$ and $\mathbf{n}(s)$ is called the osculating plane *to γ at s.*

(d) The binormal $\mathbf{b}(s)$ *to γ at s is the unit vector*

$$\mathbf{b}(s) = \mathbf{t}(s) \wedge \mathbf{n}(s).$$

The vectors $\{\mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s)\}$ form an orthonormal basis at $\gamma(s)$ which varies with s. How this basis varies with respect to s is reflected in the *Serret-Frenet formulae* (see Theorem 9). For each $s \in I$ we may write the vectors $\dot{\mathbf{t}}(s), \dot{\mathbf{n}}(s)$ and $\dot{\mathbf{b}}(s)$ in terms of the basis $\{\mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s)\}$. We already have that

$$\dot{\mathbf{t}}(s) = \kappa(s)\mathbf{n}(s) \tag{1.1}$$

and can consider similar equations for $\dot{\mathbf{n}}(s)$ and $\dot{\mathbf{b}}(s)$.

Firstly consider $\dot{\mathbf{b}}(s)$. As $\mathbf{b}(s)$ is a unit vector $\dot{\mathbf{b}}(s)$ is normal to $\mathbf{b}(s)$. Also

$$\dot{\mathbf{b}}(s) = \dot{\mathbf{t}}(s) \wedge \mathbf{n}(s) + \mathbf{t}(s) \wedge \dot{\mathbf{n}}(s) = \mathbf{t}(s) \wedge \dot{\mathbf{n}}(s).$$

Thus $\dot{\mathbf{b}}(s)$ is also normal to $\mathbf{t}(s)$ and hence must be parallel to $\mathbf{n}(s)$.

DEFINITION 8 *The* torsion $\tau(s)$ *of γ at s is the value defined by the equation*

$$\dot{\mathbf{b}}(s) = -\tau(s)\mathbf{n}(s). \tag{1.2}$$

Note that some authors, including Do Carmo, define the torsion to be $-\tau(s)$ rather than $\tau(s)$.

Now consider the components of $\dot{\mathbf{n}}(s)$ in $\mathbf{t}(s)$ and $\mathbf{b}(s)$. As $\mathbf{t}(s), \mathbf{n}(s)$ and $\mathbf{b}(s)$ are orthonormal we find

$$\begin{aligned} \dot{\mathbf{n}}(s) \cdot \mathbf{t}(s) &= -\mathbf{n}(s) \cdot \dot{\mathbf{t}}(s) &= -\kappa(s) \\ \dot{\mathbf{n}}(s) \cdot \mathbf{b}(s) &= -\mathbf{n}(s) \cdot \dot{\mathbf{b}}(s) &= \tau(s). \end{aligned}$$

Combining these equations with (1.1) and (1.2) we obtain

THEOREM 9 (Serret-Frenet) *Let $\gamma : I \to \mathbf{R}^3$ be a smooth curve parametrised by arc-length. Let $\kappa > 0$ and τ be the curvature and torsion of γ at a point s. Let $\mathbf{t}, \mathbf{n}, \mathbf{b}$ denote the tangent, normal and binormal vectors of γ at s. Then*

$$\begin{aligned} \dot{\mathbf{t}} &= & \kappa\mathbf{n} & \\ \dot{\mathbf{n}} &= -\kappa\mathbf{t} & & +\tau\mathbf{b} \\ \dot{\mathbf{b}} &= & -\tau\mathbf{n} & \end{aligned}$$

Remark 10 These formulae are often referred to simply as Frenet's formulae. They were first obtained by Frenet in a 1847 dissertation, part of which was published in 1852. Serret independently obtained the equations in 1851 – after Frenet's dissertation but before Frenet's results were widely known.

Geometrically the curvature is a measure of how quickly a curve is bending away from the tangent line; the torsion is a measure of how the curve is twisting away from the osculating plane. In particular we have the following special cases:

PROPOSITION 11 *Let $\gamma : I \to \mathbf{R}^3$ be a smooth parametrised curve with curvature κ and torsion τ.*

(a) $\kappa \equiv 0$ if and only if γ is part of a line,

(b) $\tau \equiv 0$ if and only if γ is planar,

(c) $\tau \equiv 0$ and $\kappa \equiv a^{-1} > 0$ if and only if γ is part of a circle of radius a.

PROOF:

(a) If $\kappa \equiv 0$ then $\ddot{\gamma} \equiv 0$ and so $\gamma(s) = \mathbf{u}s + \mathbf{v}$ for some vectors $\mathbf{u}, \mathbf{v} \in \mathbf{R}^3$. This is the equation of a line in $\mathbf{R}^3$.

(b) If $\tau \equiv 0$ then $\mathbf{b}$ is constant. Further

$$\frac{\mathrm{d}}{\mathrm{d}s}(\gamma \cdot \mathbf{b}) = \mathbf{t} \cdot \mathbf{b} + \gamma \cdot \dot{\mathbf{b}} = 0.$$

Hence $\gamma \cdot \mathbf{b}$ is constant which is the equation of a plane – the osculating plane in fact.

(c) From (b) we know that γ is planar. Further consider the vector $\mathbf{c} = \gamma + a\mathbf{n}$. Note from the Serret-Frenet formulae that

$$\dot{\mathbf{c}} = \mathbf{t} - a\kappa\mathbf{t} = 0$$

and hence that $\mathbf{c}$ is constant. Thus $\gamma - \mathbf{c}$ is of constant length a as required. □

PROPOSITION 12 *Let $\gamma : I \to \mathbf{R}^3$ be a smooth curve parametrised by $t \in I$ which need not be arc-length. Then the curvature κ and the torsion τ are given by*

$$\kappa = \frac{|\gamma' \wedge \gamma''|}{|\gamma'|^3} \tag{1.3}$$

and

$$\tau = \frac{(\gamma' \wedge \gamma'') \cdot \gamma'''}{|\gamma' \wedge \gamma''|^2} \tag{1.4}$$

where $'$ denotes differentiation with respect to t.

PROOF: Let s denote arc-length. From the chain rule we have that $\gamma' = s'\mathbf{t}$ and that

$$\gamma'' = s''\mathbf{t} + s'\mathbf{t}' = s''\mathbf{t} + (s')^2\kappa\mathbf{n}.$$

Thus

$$\gamma' \wedge \gamma'' = (s')^3\kappa\mathbf{b} = |\gamma'|^3\kappa\mathbf{b}$$

and equation (1.3) follows. Now

$$\gamma''' \cdot \mathbf{b} = (s')^2\kappa\mathbf{n}' \cdot \mathbf{b} = (s')^3\kappa\tau.$$

Hence

$$(\gamma' \wedge \gamma'') \cdot \gamma''' = |\gamma'|^6\kappa^2\tau$$

and substituting in equation (1.3) we obtain the required expression for τ. □

The next theorem is not on the a2 geometry syllabus and will not be proved. However knowing the theorem justifies further the notions of curvature and torsion. A proof can be found in Do Carmo pp. 309-311 but relies more on the theory of differential equations rather than on geometry.

THEOREM 13 *(*) :(The Fundamental Theorem of the Local Theory of Curves).*
Given differential functions $\kappa(s) > 0$ and $\tau(s)$, $s \in I$, there exists a smooth parametrised curve $\gamma : I \to \mathbf{R}^3$ such that s is the arc-length, κ is the curvature and τ is the torsion of γ. Further, a second curve α satisfying the same conditions differs from γ only by a rigid motion of $\mathbf{R}^3$.

So the curvature and torsion of a curve in $\mathbf{R}^3$ determine the curve save for its position in space. This really should be no surprise. Changes in the Serret-Frenet trihedron $\{\mathbf{t}, \mathbf{n}, \mathbf{b}\}$ of a curve are decided by curvature and torsion through the Serret-Frenet formulae. Given some initial conditions for this trihedron and functions κ, τ which dictate how it varies, then $\{\mathbf{t}, \mathbf{n}, \mathbf{b}\}$, and hence the curve are uniquely determined.

The theory of space curves generalises to curves in $\mathbf{R}^n$. In these cases there are $n-1$ curvature functions, a moving Serret-Frenet basis containing n vectors and relations similar to the Serret-Frenet formulae above determining changes in this basis. Again given functions $\kappa_1, ..., \kappa_{n-1}$ with $\kappa_1, ..., \kappa_{n-2} > 0$ there is a unique curve with curvature functions $\kappa_1, ..., \kappa_{n-1}$. For $n = 3$, $\kappa_1 = \kappa$ and $\kappa_2 = \tau$. (See Klingenberg *A Course in Differential Geometry*, pp. 13-14).

For a brief survey of certain aspects of the global theory of curves, there really is no better introduction than Do Carmo §1.7.

Example 14 *Q: Find the curvature and torsion at each point of the curve given parametrically by*

$$\mathbf{r}(t) = (2t, t^2, \frac{1}{3}t^3).$$

A: We have

$$\mathbf{r}'(t) = (2, 2t, t^2), \qquad \mathbf{r}''(t) = (0, 2, 2t), \qquad \mathbf{r}'''(t) = (0, 0, 2).$$

Thus

$$\begin{aligned} |\mathbf{r}'(t)| &= t^2 + 2, \\ |\mathbf{r}'(t) \wedge \mathbf{r}''(t)| &= |(2t^2, -4t, 4)| = 2(t^2 + 2), \\ (\mathbf{r}'(t) \wedge \mathbf{r}''(t)) \cdot \mathbf{r}'''(t) &= 8. \end{aligned}$$

Substituting these values into equations (1.3) and (1.4) we obtain

$$\kappa(t) = \frac{2}{(t^2+2)^2}, \qquad \tau(t) = \frac{2}{(t^2+2)^2}. \qquad \square$$

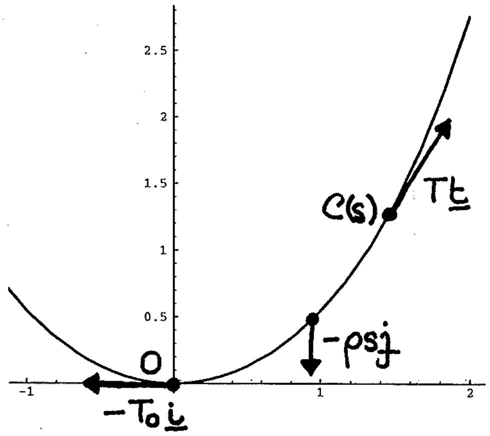


Figure 3 ($k = 1$)

The forces acting on the segment of chain between (0,0) and $C(s)$ are:

the weight of the chain $= -\rho s\mathbf{j}$,
the tension at $C(s)$ $= T(s)\mathbf{t}(s)$,
the tension at (0,0) $= -T_0\mathbf{i}$.

Example 15 *Q: An infinite uniform chain C hangs under its own weight in the xy-plane with lowest point* $(0,0)$*. The chain has density ρ and the tension at* $(0,0)$ *is T_0. Find the equation of the chain.*

A: Parametrise the curve $s \mapsto C(s)$ by arc-length from $(0,0)$. We shall denote the angle between the curve at $C(s)$ and the x-axis by $\psi(s)$ and the tension in the chain at $C(s)$ by $T(s)$. The forces acting on the length of chain between (0,0) and $C(s)$ (see Figure 3) are in equilibrium giving:

$$T(s)\mathbf{t}(s) = T_0\mathbf{i} + \rho s\mathbf{j}.$$

Taking components in the direction of $\mathbf{n}(s) = (-\sin\psi, \cos\psi)$ we find

$$T_0 \sin\psi = \rho s \cos\psi.$$

Hence

$$s = k \tan\psi \text{ where } k = T_0/\rho. \qquad \square \tag{1.5}$$

Remark 16 Equation (1.5) is the equation of the chain in the co-ordinates (s, ψ) which are known as *intrinsic co-ordinates.* We can find the equation of C in Cartesian co-ordinates using the Fundamental Theorem and a little inspection.

In intrinsic co-ordinates the tangent to a given curve equals $\mathbf{t} = (\cos\psi, \sin\psi)$ and so

$$\kappa\mathbf{n} = \frac{d\mathbf{t}}{ds} = (-\sin\psi, \cos\psi)\frac{d\psi}{ds}.$$

Hence

$$\kappa(s) = \frac{\mathrm{d}\psi}{\mathrm{d}s}.$$

So the curvature of the chain C is given by

$$\kappa(s) = \frac{\mathrm{d}}{\mathrm{d}s}\left(\tan^{-1}\left(\frac{s}{k}\right)\right) = \frac{k}{s^2+k^2}.$$

By the Fundamental Theorem any curve in the xy-plane with this curvature and a minimum at (0,0) must coincide with the chain. This is left as an exercise for the reader.

Exercise 17 Show that the curve

$$y = k\left(\cosh\left(\frac{x}{k}\right) - 1\right)$$

has curvature $\kappa(s) = k/(s^2+k^2)$, where s is arc-length, and a minimum at (0,0). This curve is known as a *catenary*, from the Latin word 'catena' meaning 'chain'.

1.3 Exercises

Exercise 18 An epicycloid is obtained as the locus of a point on the circumference of a circle of radius r which rolls without slipping on a circle of the same radius. Find the length of the epicycloid.

Exercise 19 Find the curvature and torsion of the *elliptical helix*

$$(a\cos t, b\sin t, ct) \qquad a,b,c>0,\ t\in\mathbf{R}.$$

Exercise 20 Calculate $\kappa, \tau, \mathbf{t}, \mathbf{n}$ and $\mathbf{b}$ for the curve

$$\left(\frac{1}{3}(1+s)^{3/2}, \frac{1}{3}(1-s)^{3/2}, \frac{s}{\sqrt{2}}\right), \qquad -1<s<1.$$

Exercise 21 Let $\gamma(s)$ be a curve parametrised by arc-length with nowhere vanishing curvature κ and torsion τ. Show that γ lies on a sphere if and only if

$$\frac{\tau}{\kappa} = \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{\tau\kappa^2}\frac{\mathrm{d}\kappa}{\mathrm{d}s}\right).$$

Exercise 22 Let $\gamma : [a,b] \to S$ be a smooth closed curve parametrised by arc-length with torsion $\tau(s)$ lying on a sphere S. (That is: $\gamma|_{(a,b)} : (a,b) \to S$ is a smooth parametrised curve with $\gamma(a) = \gamma(b)$ and $\gamma'(a) = \gamma'(b)$.) Show that $\int_a^b \tau(s)\,\mathrm{d}s$ is zero.

Exercise 23 A smooth curve γ is called a *helix* if the unit tangent $\mathbf{t}$ has constant angle θ with some fixed unit vector $\mathbf{u}$. i.e. $\mathbf{t}\cdot\mathbf{u} = \cos\theta$ at each point of the curve.

Let γ be a smooth curve with nowhere vanishing curvature. Show that γ is a helix if and only if τ/κ is constant. [Hint: Consider the vector

$$\mathbf{u} = \cos\theta\mathbf{t} + \sin\theta\mathbf{b}$$

where θ satisfies $\cot\theta = \tau/\kappa$.]

Exercise 24 Show that the curve

$$\gamma_1(t) = (t+\sqrt{3}\sin t, 2\cos t, \sqrt{3}t - \sin t)$$

is a helix by computing its curvature and torsion. Find a helix γ_2 of the form $(a\cos t, a\sin t, bt)$ and a rigid motion T of $\mathbf{R}^3$ such that $T\gamma_1 = \gamma_2$.

Exercise 25 Find two rigid motions of $\mathbf{R}^3$ carrying the parabola $(\sqrt{2}t, t^2, 0)$ to the parabola $(-t, t, t^2)$.

Chapter 2

Surfaces

2.1 Parametrised Surfaces

Just as the curves we have studied are images of the real line twisted and stretched into new shapes in Euclidean space, then surfaces are essentially twisted and stretched images of $\mathbf{R}^2$. We defined a curve to be given by a function $\mathbf{r} : \mathbf{R} \to \mathbf{R}^3$. Only one parametrisation was needed and such curves are all homeomorphic to the real line. Without mentioning it, we omitted curves such as the circle and ellipse and other closed curves.

Within the limits of this course we shall discuss 'parametrised surfaces' and I shall take this to mean a surface with a single given parametrisation. We will again have to omit such examples as spheres, tori and cylinders say. There is no hope of parametrising a sphere or a torus with a single set of co-ordinates because these (compact) surfaces are not homeomorphic to the (non-compact) plane. More generally a surface is defined as a space that looks locally like $\mathbf{R}^2$ and is covered with several, possibly infinitely many co-ordinate systems. This aspect of surface theory is a little beyond the scope of the course and is covered in more detail in §2.5.

This restriction to parametrised surfaces is a small one as we are studying local properties of surfaces. Also while surfaces such as tori and spheres cannot be parametrised entirely by a single parametrisation, we may parametrise open dense subsets and so we can calculate certain global properties (e.g. area) of these surfaces.

DEFINITION 26 *A* smooth parametrised surface *is a map, known as a* chart,

$$\mathbf{r} : U \to \mathbf{R}^3 : (u, v) \to (x(u, v), y(u, v), z(u, v))$$

from an open subset $U \subseteq \mathbf{R}^2$ to $\mathbf{R}^3$ such that

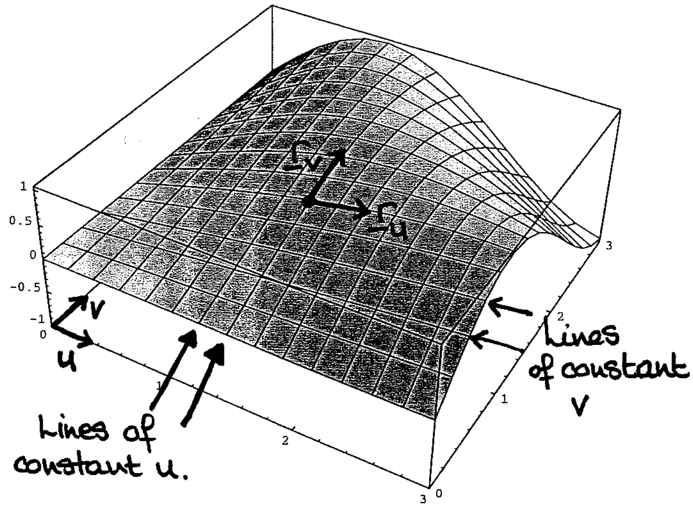
- $\mathbf{r}$ *is smooth i.e. x, y, z have continuous partial derivatives of all orders,*
- $\mathbf{r} : U \to \mathbf{r}(U)$ *is a homeomorphism,*
- *at each point of $\mathbf{r}(U)$ the vectors*

$$\mathbf{r}_u = \frac{\partial \mathbf{r}}{\partial u} \text{ and } \mathbf{r}_v = \frac{\partial \mathbf{r}}{\partial v}$$

  are linearly independent.

Our definition above of a parametrised surface differs with that of Do Carmo (p.78) in that we require $\mathbf{r}$ to be a homeomorphism onto its image. Without this requirement then self-intersections may occur, whereas a parametrised surface (by the above definition) is a genuine example of a smooth surface (Do Carmo p.52). The difference between the two definitions is only slight – without the requirement that $\mathbf{r} : U \to \mathbf{r}(U)$ is a homeomorphism, it is still true that around each $p \in U$ there is a neighbourhood $V \subset U$ such that $\mathbf{r} : V \to \mathbf{r}(V)$ is a smooth parametrised surface (see Do Carmo p.79).

Figure 4
The Graph of $\sin(\frac{1}{2}uv)$
$(0 \leq u, v \leq 3)$



DEFINITION 27 *Let* $\mathbf{r}(U)$ *be a smooth parametrised surface and let* $p = \mathbf{r}(u_0, v_0)$*. The* tangent plane *to* $\mathbf{r}(U)$ *at* p *is the vector space spanned by the vectors*

$$\mathbf{r}_u(u_0, v_0) \text{ and } \mathbf{r}_v(u_0, v_0).$$

The tangent plane to $\mathbf{r}(U)$ at p is denoted by $T_p(\mathbf{r}(U))$ and elements of $T_p(\mathbf{r}(U))$ are called *tangent vectors* to $\mathbf{r}(U)$ at p. It is easy to check that the tangent plane at p is the plane spanned by the tangent vectors to all curves in $\mathbf{r}(U)$ which pass through p (see Do Carmo p.83).

DEFINITION 28 *Let* $\mathbf{r}(U)$ *be a smooth parametrised surface in* $\mathbf{R}^3$*. A* normal vector *to* $\mathbf{r}(U)$ *at the point* p *is any (non-zero) vector orthogonal to* $T_p(\mathbf{r}(U))$*.*

The normal vectors are non-zero scalar multiples of $\mathbf{r}_u \wedge \mathbf{r}_v$ where $\wedge$ denotes the cross or vector product in $\mathbf{R}^3$. The two unit vectors

$$\pm \frac{\mathbf{r}_u \wedge \mathbf{r}_v}{|\mathbf{r}_u \wedge \mathbf{r}_v|}$$

are known as the *unit normal vectors* to $\mathbf{r}(U)$ at p.

Note that a parametrised surface is a surface in $\mathbf{R}^3$ with a preferred choice of co-ordinates. The image $X = \mathbf{r}(U)$ in $\mathbf{R}^3$ of a given chart is also the image of other charts. Each such chart gives a different parametrised surface but we would hope that any questions asked of X (simply as a subspace of $\mathbf{R}^3$) such as, "what is the area of X?" and "what is the length of a curve in X?", will yield the same answers, irrespective of what chart we use. This will be an important consideration in all future definitions – that any new definitions are chart independent.

Before we discuss any of the theory of surfaces it seems best to introduce some of the more important examples. Due to the nature of this course I have kept the following examples simple and brief, but I hope they prove sufficiently various to interest the reader. Further examples can be found in any standard text on surfaces.

Example 29 *The Sphere*

Consider the map $\mathbf{r}_1 : (-\pi, \pi) \times (0, \pi) \to \mathbf{R}^3$ (see Figure 5) given by

$$\mathbf{r}_1 : (u, v) \mapsto (\cos u \sin v, \sin u \sin v, \cos v).$$

It is easy to check that the image of this map is contained in S^2, the unit sphere centred at the origin. In fact the image is the whole sphere save for half a great circle. The parameter u is the angle between the projection of $\mathbf{r}_1(u, v)$ onto the xy-plane and the x-axis and v is the angle between $\mathbf{r}_1(u, v)$ and the z-axis.

Consider also the map $\mathbf{r}_2 : \mathbf{R}^2 \to \mathbf{R}^3$ defined by

$$\mathbf{r}_2 : (u, v) \mapsto \left(\frac{2u}{u^2 + v^2 + 1}, \frac{2v}{u^2 + v^2 + 1}, \frac{1 - u^2 - v^2}{u^2 + v^2 + 1} \right).$$

This again is a chart of the unit sphere. The map $\mathbf{r}_2$ is in fact stereographic projection (see Figure 6) from the 'south pole' $S = (0, 0, -1)$; that is a point of $(u, v) \in \mathbf{R}^2$ is mapped to the intersection of the sphere with the line joining (u, v) and S. In this case the image of the sphere is the whole sphere minus S.

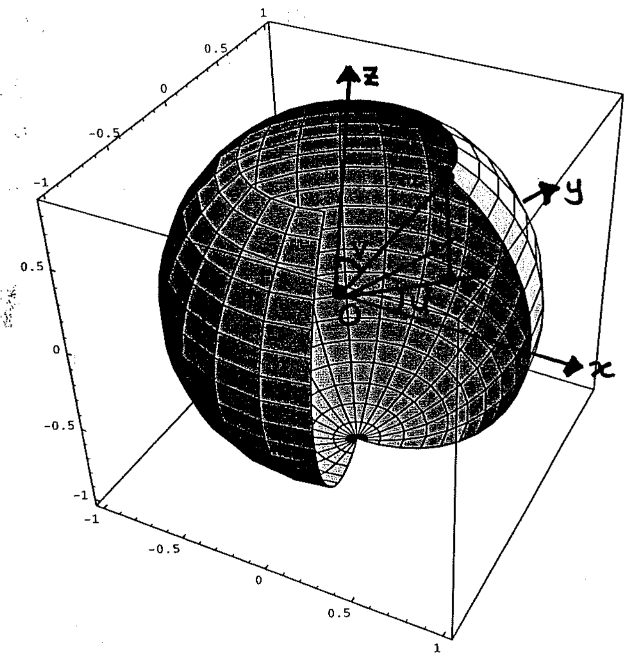


Figure 5 – Spherical Co-ordinates

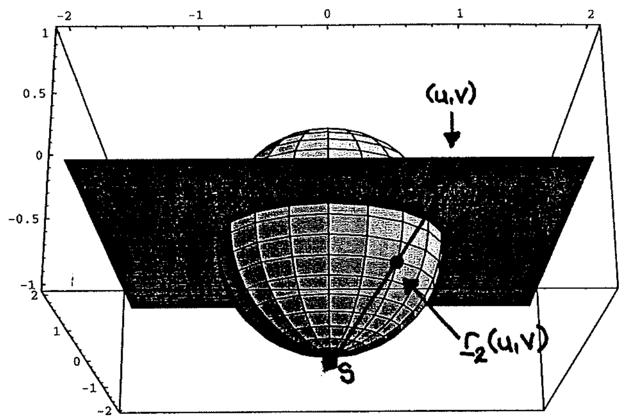


Figure 6 – Stereographic Projection

Example 30 *Graphs*

Amongst the simplest examples of parametrised surfaces are graphs. Let $f(x, y)$ be a smooth function defined on an open set $U \subseteq \mathbf{R}^2$. Then the *graph* of f is the surface $z = f(x, y)$ and may be parametrised by

$$\mathbf{r}(u, v) = (u, v, f(u, v)) \qquad (u, v) \in U.$$

These graphs seem almost too simple a family of surfaces to be of interest. One point of importance though is that any smooth surface in $\mathbf{R}^3$ is, locally at least, a graph. That is (Do Carmo p.63):

- About any point of a smooth surface there is an open neighbourhood U such that U is a graph of the form $z = f(x, y)$ or $y = f(x, z)$ or $x = f(y, z)$ for some smooth function f.

Exercise 31 Express the cylinder $x^2 + y^2 = a^2$ for $a \neq 0$ as the union of parametrised surfaces which are graphs.

Example 32 *The Cone*

The punctured cone $x^2 + y^2 = z^2$, $(z > 0)$ in $\mathbf{R}^3$ may be smoothly parametrised by

$$\mathbf{r}(u, v) = (u, v, \sqrt{u^2 + v^2}) \qquad u, v \in \mathbf{R}, u^2 + v^2 \neq 0.$$

Note that the two sheeted cone $x^2+y^2 = z^2$ is not the image of any chart as no neighbourhood of the cone about $(0, 0, 0)$ is homeomorphic to an open subset of $\mathbf{R}^2$. (Why is this?)

Consider now the one sheeted cone C given by $x^2 + y^2 = z^2$, $(z \geq 0)$. This certainly is the image of a chart $\mathbf{s} : \mathbf{R}^2 \to C$, but for no such chart is C smooth at the point $(0, 0, 0)$. To prove this we assume that the cone may be locally parametrised about $(0, 0, 0)$ as the graph of a smooth function. The only possibility (from $z = f(x, y)$ or $y = f(x, z)$ or $x = f(y, z)$) is a graph of the form $z = f(x, y)$ and by the definition of C we see that

$$f(x, y) = \sqrt{x^2 + y^2}.$$

As f is not differentiable at $(0, 0)$ then $(0, 0, 0)$ is not a smooth point of C for *any* parametrisation. Such points on a surface are called *singular points*.

Example 33 *Surfaces of Revolution*

Surfaces may also be formed by taking a curve in $\mathbf{R}^3$ and using this curve to generate a surface. Two such families are *surfaces of revolution* and (below) *ruled surfaces.* A surface of revolution is formed by rotating a smooth curve in (say) the xz-plane about the z-axis. For example the cylinder in the above exercise is a surface of revolution.

Assume the curve has equation $x = f(z)$. Then the surface of revolution generated has equation $x^2 + y^2 = f(z)^2$. The surface cannot entirely be parametrised with one co-ordinate system but the map

$$\mathbf{r}(\theta, z) = (f(z)\cos\theta, f(z)\sin\theta, z) \qquad \theta \in (0, 2\pi), z \in \mathbf{R}$$

parametrises all of the surface except for the original generating curve. The curves of the form $\theta =$ const. are called *meridians* and those with equations $z =$ const. are *parallels.*

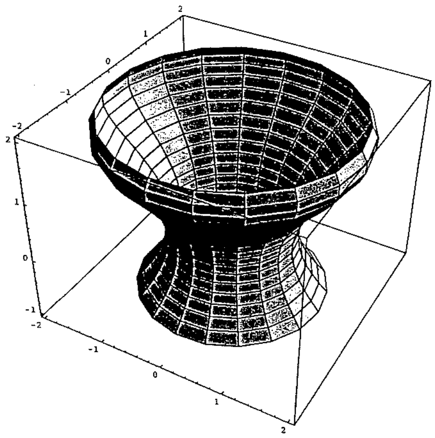


Figure 7 – A Surface of Revolution

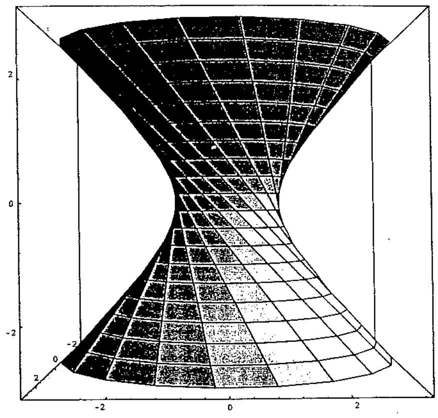


Figure 8 – A Hyperboloid of One Sheet

Example 34 *Ruled Surfaces*

Let $\gamma : I \to \mathbf{R}^3$ be a smooth curve in $\mathbf{R}^3$ and let $\mathbf{w} : I \to \mathbf{R}^3 - \{\mathbf{0}\}$ be a second non-vanishing vector function on I. Then the parametrised surface given by

$$\mathbf{r}(u, v) = \gamma(u) + v\mathbf{w}(u) \quad u \in I, v \in \mathbf{R}$$

is an example of a *ruled surface.* The curve γ is known as the *directrix* and the lines in the surface given by $u =$ constant are known as *rulings.*

Note that the chart $\mathbf{r}$ above need not be a homeomorphism onto its image and so such a ruled surface may have self-intersections, although these may be avoided by limiting the domain of the co-ordinate v. For example the image of the map

$$\mathbf{r}(u, v) = (v\cos u, v\sin u, v) \qquad u \in (0, 2\pi), v \in \mathbf{R},$$

is all of the two sheeted cone except for two rays (two halves of the line $x = z$). The map $\mathbf{r}$ is not a chart as $(0, 0, 0)$ is a self-intersection. However the restriction of $\mathbf{r}$ to $(0, 2\pi) \times (0, \infty)$ is a chart for the one sheeted cone except for a single ruling.

Exercise 35 Show that the hyperbolic paraboloid $z = xy$ and the hyperboloid of one sheet $x^2 + y^2 = z^2 + 1$ in $\mathbf{R}^3$ are ruled surfaces.

To conclude this section we briefly discuss smooth maps between parametrised surfaces. We say that a map $f : U \to \mathbf{R}^2$, from an open subset $U \subseteq \mathbf{R}^2$ is *smooth* on U if all partial derivatives of f of all orders exist at all points of U. As a parametrised surface is a space endowed with co-ordinates from $\mathbf{R}^2$ then we may extend this to the idea of smooth maps between parametrised surfaces.

DEFINITION 36 *Let* $\mathbf{r} : U \to \mathbf{R}^3$ *and* $\mathbf{s} : V \to \mathbf{R}^3$ *be smooth parametrised surfaces in* $\mathbf{R}^3$. *We say that a map* $f : \mathbf{r}(U) \to \mathbf{s}(V)$ *is* smooth *if the map*

$$\mathbf{s}^{-1} \circ f \circ \mathbf{r} : U \to V$$

has partial derivatives of all orders. A map f *between parametrised surfaces which is smooth and has a smooth inverse is known as a* diffeomorphism.

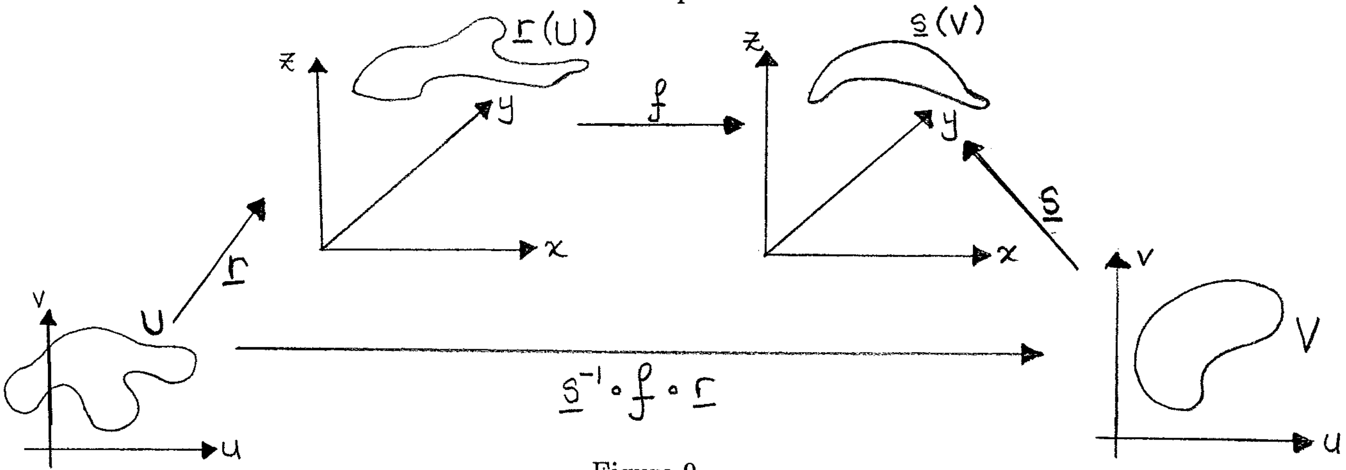


Figure 9

Put into this context a parametrised surface (as given in Definition 26) is a subspace of $\mathbf{R}^3$ which is diffeomorphic to $\mathbf{R}^2$ and we see that all parametrised surfaces are diffeomorphic to one another. Therefore the notion of a diffeomorphism is a rather limited concept when dealing simply with smooth parametrised surfaces – from a differential geometric point of view, parametrised surfaces are all equivalent, although this is far from true in the larger class of smooth surfaces (see §2.5). For most of this course we will be interested in the local intrinsic geometry of surfaces i.e. the metric properties of the surface.

Exercise 37 Let X denote the paraboloid $z = x^2 + y^2$ in $\mathbf{R}^3$. Give a parametrisation for X and prove, for this parametrised surface, that rotations of X about the z-axis are diffeomorphisms.

2.2 The First Fundamental Form: Lengths, Areas and Isometries

Let $U \subseteq \mathbf{R}^2$ be an open subset of the plane and $\mathbf{r} : U \to \mathbf{R}^3$ be a chart of a smooth surface X. Let

$$\gamma : I \to X : \gamma(t) = \mathbf{r}(u(t), v(t))$$

be a smooth curve lying in X.

DEFINITION 38 *We define the* length *of* γ *to be*

$$\mathcal{L}(\gamma) = \int_I \left| \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right| \mathrm{d}t. \tag{2.1}$$

Using the chain rule it is easy to see that the length of γ does not depend on the choice of parameter t.

Now

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \frac{\mathrm{d}u}{\mathrm{d}t}\frac{\partial \mathbf{r}}{\partial u} + \frac{\mathrm{d}v}{\mathrm{d}t}\frac{\partial \mathbf{r}}{\partial v}$$

or written more concisely

$$\dot{\gamma} = \dot{u}\mathbf{r}_u + \dot{v}\mathbf{r}_v.$$

So the length of γ equals

$$\int_I \sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}\, \mathrm{d}t \tag{2.2}$$

where

$$E = \mathbf{r}_u \cdot \mathbf{r}_u, \quad F = \mathbf{r}_u \cdot \mathbf{r}_v, \quad G = \mathbf{r}_v \cdot \mathbf{r}_v.$$

DEFINITION 39 *The quadratic form*

$$\alpha \mathbf{r}_u + \beta \mathbf{r}_v \mapsto E\alpha^2 + 2F\alpha\beta + G\beta^2$$

on the tangent plane T_pX *is called the* first fundamental form *of* X.

What does this actually mean? The first fundamental form is the quadratic form on the tangent plane to X at the point $\mathbf{r}(u,v)$ given by

$$\mathbf{x} \mapsto |\mathbf{x}|^2$$

where $\mathbf{x}$ is a tangent vector to the surface. Now $\{\mathbf{r}_u, \mathbf{r}_v\}$ is a basis for the tangent plane and with respect to this basis the first fundamental form has coefficients $E, 2F$ and G. Geometrically it can be thought of as the square of the element of arc-length and one often writes

$$\mathrm{d}s^2 = E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + G\mathrm{d}v^2.$$

Precisely what '$\mathrm{d}u$' and '$\mathrm{d}v$' mean is explained in further detail in §2.5 for interested readers. However one thinks about the first fundamental form, remember that the form is associated with the surface. If one changes parameters this element of arc-length will not change, nor the quadratic form on the tangent plane; however the expression for the first fundamental form will generally look different in terms of the new variables.

Example 40 *Q: Find the first fundamental form of the plane using (i) Cartesian co-ordinates and (ii) polar co-ordinates.*

A: Using Cartesian co-ordinates we find

$$\mathbf{r}(u,v) = (u,v), \qquad u,v \in \mathbf{R}$$

and with polar co-ordinates

$$\mathbf{R}(r,\theta) = (r\cos\theta, r\sin\theta), \qquad r > 0, \theta \in (0, 2\pi).$$

So

$$\begin{aligned} \mathbf{r}_u &= (1,0), & \mathbf{r}_v &= (0,1), \\ \mathbf{R}_r &= (\cos\theta, \sin\theta), & \mathbf{R}_\theta &= (-r\sin\theta, r\cos\theta). \end{aligned}$$

With respect to the two co-ordinate systems the first fundamental form is:

$$\mathrm{d}u^2 + \mathrm{d}v^2 \qquad \text{and} \qquad \mathrm{d}r^2 + r^2\mathrm{d}\theta^2. \qquad \square$$

How might one now calculate the area of a region of X? Let $V \subset U$ be an open subset of U; we wish to calculate the area of $\mathbf{r}(V)$. Consider a small parallelogram with vertices

$$\mathbf{r}(u,v), \quad \mathbf{r}(u+\delta u, v), \quad \mathbf{r}(u, v+\delta v), \quad \mathbf{r}(u+\delta u, v+\delta v).$$

Now

$$\mathbf{r}(u+\delta u, v) - \mathbf{r}(u,v) = \mathbf{r}_u(u,v)\delta u + O(\delta u^2)$$

and there is a similar expression for varying v. So the area of the parallelogram is, ignoring higher order terms,

$$|\mathbf{r}_u \wedge \mathbf{r}_v|\, \delta u\, \delta v.$$

It thus seems reasonable to define:

DEFINITION 41 *The* area *of* $\mathbf{r}(V)$ *equals*

$$\int_V |\mathbf{r}_u \wedge \mathbf{r}_v|\, \mathrm{d}u\, \mathrm{d}v. \tag{2.3}$$

Now

$$\begin{aligned}|\mathbf{r}_u \wedge \mathbf{r}_v|^2 &= (\mathbf{r}_u \wedge \mathbf{r}_v)\cdot(\mathbf{r}_u \wedge \mathbf{r}_v)\\ &= (\mathbf{r}_u\cdot\mathbf{r}_u)(\mathbf{r}_v\cdot\mathbf{r}_v) - (\mathbf{r}_u\cdot\mathbf{r}_v)(\mathbf{r}_v\cdot\mathbf{r}_u)\\ &= EG - F^2.\end{aligned}$$

Thus the expression (2.3) for the area of $\mathbf{r}(V)$ can be rewritten as

$$\int_V \sqrt{EG-F^2}\,\mathrm{d}u\,\mathrm{d}v. \tag{2.4}$$

Exercise 42 Show that the definition of area given above is independent of the choice of parametrisation. (You will need to use the change of variables formula for double integrals – see Apostol, *Mathematical Analysis* p.421.)

Example 43 *Q: Show that the area of a sphere of radius a equals* $4\pi a^2$.

A: We may parametrise the sphere by setting

$$\mathbf{r}(u,v) = (a\cos u\sin v, a\sin u\sin v, a\cos v) \qquad u\in(-\pi,\pi),\ v\in(0,\pi),$$

omitting only half a great circle. Then

$$\begin{aligned}\mathbf{r}_u &= (-a\sin u\sin v, a\cos u\sin v, 0),\\ \mathbf{r}_v &= (a\cos u\cos v, a\sin u\cos v, -a\sin v).\end{aligned}$$

Thus (with respect to the co-ordinates u and v) the first fundamental form is given by

$$E = a^2\sin^2 v, \quad F = 0, \quad G = a^2$$

and the area is given by

$$\int_0^\pi\int_{-\pi}^{\pi} a^2|\sin v|\,\mathrm{d}u\,\mathrm{d}v = 2\pi a^2\int_0^\pi \sin v\,\mathrm{d}v = 4\pi a^2$$

as required. □

Example 44 *Q: The* tractoid *(see Figure 12) is the surface of revolution formed by rotating the curve*

$$x(t) = -(\cos t + \log\tan\frac{t}{2}), \qquad y(t) = \sin t, \qquad t\in(0,\pi/2)$$

(known as the tractrix*) about the x-axis. Show that when the tractrix is parametrised by arc-length s the first fundamental form of the tractoid is*

$$\mathrm{d}s^2 + e^{-2s}\mathrm{d}\theta^2. \tag{2.5}$$

Show that the area of the tractoid equals 2π.

A: We may parametrise the tractoid by writing

$$\mathbf{r}(t,\theta) = (x(t), y(t)\cos\theta, y(t)\sin\theta), \qquad t\in(0,\infty),\ \theta\in(0,2\pi),$$

omitting only the original tractrix. Differentiating with respect to t and θ we find that

$$\begin{aligned}\mathbf{r}_t &= (-\cos t\cot t, \cos t\cos\theta, \cos t\sin\theta),\\ \mathbf{r}_\theta &= (0, -\sin t\sin\theta, \sin t\cos\theta).\end{aligned}$$

Thus the first fundamental form is given by

$$\cot^2 t\,\mathrm{d}t^2 + \sin^2 t\,\mathrm{d}\theta^2. \tag{2.6}$$

Now

$$\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 = \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 = \left(\frac{\cos^2 t}{\sin t}\right)^2 + \cos^2 t = \cot^2 t.$$

As s is decreasing with respect to t then $\mathrm{d}s/\mathrm{d}t = -\cot t$ and hence $s = -\log\sin t$. Substituting these expressions into (2.6) we obtain (2.5).

The area of the tractoid is then given by the integral

$$\int_0^\infty \int_0^{2\pi} e^{-s}\,\mathrm{d}\theta\,\mathrm{d}s = 2\pi. \qquad \square$$

Properties of surfaces which depend solely on the first fundamental such as length and area (and geodesics and Gaussian curvature – see later) are called *intrinsic.* Maps between surfaces which preserve the intrinsic geometry are called *isometries.*

DEFINITION 45 *An* isometry *between two surfaces X and Y is a diffeomorphism $f: X \to Y$ which maps curves in X to curves in Y of the same length. X and Y are then said to be* isometric.

As the first fundamental form represents an element of arc-length then the following theorem should be intuitively clear.

THEOREM 46 *Two surfaces X and Y are isometric if and only if there exist an open subset $U \subset \mathbf{R}^2$ and parametrisations*

$$\mathbf{r}: U \to X, \qquad \mathbf{s}: U \to Y,$$

such that the first fundamental forms of X and Y are the same.

PROOF: Sufficiency is easy. Suppose two such parametrisations $\mathbf{r}$ and $\mathbf{s}$ exist with the same fundamental forms – I claim $f = \mathbf{s}\mathbf{r}^{-1} : X \to Y$ is the required isometry. Let C be a smooth curve in U. The lengths of $\mathbf{r}(C)$ and $\mathbf{s}(C) = f(\mathbf{r}(C))$ are identical as they are given by the same integral (2.2).

Suppose now that $f: X \to Y$ is an isometry of two surfaces and suppose that $\mathbf{r}: U \to X$ is a parametrisation of X. Let $\mathbf{s} = f\mathbf{r}: U \to Y$. We shall write $E, 2F, G$ and $\tilde{E}, 2\tilde{F}, \tilde{G}$ for the coefficients of X and Y with respect to $\mathbf{r}$ and $\mathbf{s}$. As f is an isometry we have that

$$\int_I \sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}\,\mathrm{d}t = \int_I \sqrt{\tilde{E}\dot{u}^2 + 2\tilde{F}\dot{u}\dot{v} + \tilde{G}\dot{v}^2}\,\mathrm{d}t \tag{2.7}$$

for all smooth curves $(u(t), v(t))$, $t \in I$ in U. Firstly we choose part of a co-ordinate curve, namely: $u(t) = u_0 + t$ and $v(t) = v_0$ for $t \in (0, \epsilon)$ and for some point $(u_0, v_0) \in U$. By the continuity of E and $\tilde{E}$ and applying (2.7) above we find

$$\begin{aligned}\sqrt{E(u_0, v_0)} &= \lim_{\epsilon\to 0} \frac{1}{\epsilon}\int_0^\epsilon \sqrt{E(u_0 + t, v_0)}\,\mathrm{d}t \\ &= \lim_{\epsilon\to 0} \frac{1}{\epsilon}\int_0^\epsilon \sqrt{\tilde{E}(u_0 + t, v_0)}\,\mathrm{d}t = \sqrt{\tilde{E}(u_0, v_0)},\end{aligned}$$

and hence $E = \tilde{E}$. By similar arguments using the curves

$$\begin{aligned} u(t) &= u_0, & v(t) &= v_0 + t, & t &\in (0, \epsilon), \\ u(t) &= u_0 + t, & v(t) &= v_0 + t, & t &\in (0, \epsilon), \end{aligned}$$

we may conclude that $G = \tilde{G}$ and that $F = \tilde{F}$. $\square$

Example 47 *Q: The* catenoid *(with a meridian removed) and* helicoid *are respectively parametrised by*

$$\begin{aligned} \mathbf{r}(u, v) &= (u, \cosh u \cos v, \cosh u \sin v), & & u \in \mathbf{R}, v \in (0, 2\pi), \\ \mathbf{s}(\tilde{u}, \tilde{v}) &= (\tilde{u}, \tilde{v}\cos\tilde{u}, \tilde{v}\sin\tilde{u}), & & \tilde{u} \in \mathbf{R}, \tilde{v} \in \mathbf{R}. \end{aligned}$$

Show that the catenoid is isometric to part of the helicoid, in such a way that meridians of the catenoid map to rulings of the helicoid.

A: The first fundamental form of the catenoid equals

$$\cosh^2 u\, du^2 + \cosh^2 u\, dv^2$$

and the first fundamental form of the helicoid equals

$$(1 + \tilde{v}^2)\, d\tilde{u}^2 + d\tilde{v}^2. \tag{2.8}$$

Now consider the map

$$\mathbf{r}(u, v) \mapsto \mathbf{s}(v, \sinh u) \text{ for } u \in \mathbf{R}, v \in (0, 2\pi) \tag{2.9}$$

between the catenoid and the helicoid. Under the substitution $\tilde{u} = v$ and $\tilde{v} = \sinh u$ then the form (2.8) becomes

$$(1 + \sinh^2 u)\, du^2 + d(\sinh u)^2 = \cosh^2 u\, du^2 + \cosh^2 u\, dv^2$$

which is the first fundamental form of the catenoid. Thus the map (2.9) is indeed an isometry.

The meridians of the catenoid are given by the equations $v =$ *constant*. Under the above isometry the meridians map to the curves on the helicoid given by $\tilde{u} =$ *constant* – i.e. the rulings.

□

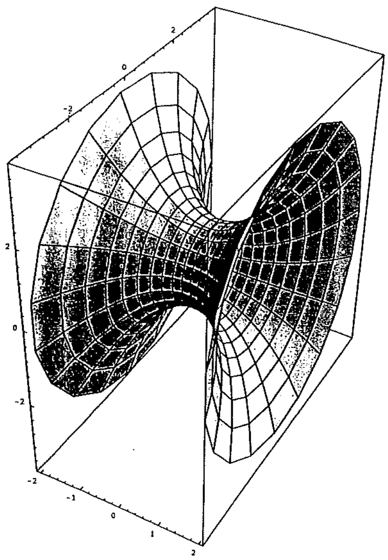


Figure 10 – The Catenoid

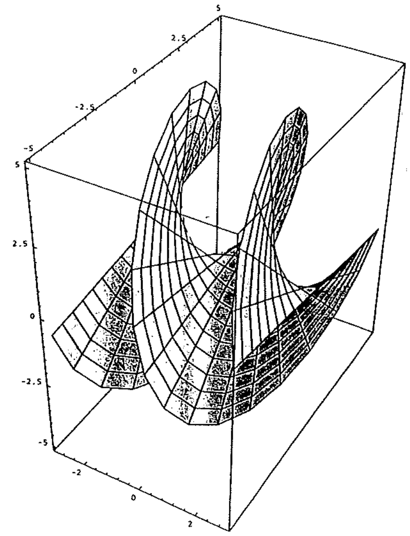


Figure 11 – The Helicoid

Remark 48 (*) What follows in the remainder of this section is not on the syllabus of the a2 course. It is an attempt to motivate the definitions for length and area (2.2), (2.4) we have introduced above and also a digression on abstract geometric surfaces. Whilst the discussion is not self-contained those parts of the notes that refer back to it likewise carry an asterisk.

Thus far we have not made any calculations of lengths and areas which couldn't have been done as easily with the old expressions (2.1), (2.3) as with the new expressions (2.2), (2.4) which are in terms of coefficients of the first fundamental form. The calculations in the following examples however can only be done using the new definitions of length and area.

Example 49 *Q: The* flat torus $\mathbf{T}$ *is the surface in* $\mathbf{R}^4$ *given by*

$$\mathbf{T} = \{(x, y, z, t) : x^2 + y^2 = z^2 + t^2 = 1\}.$$

Show that $\mathbf{T}$ *is locally isometric to* $\mathbf{R}^2$ *and calculate the area of* $\mathbf{T}$.

A: We may parametrise (a dense open subset of) $\mathbf{T}$ by the chart

$$\mathbf{r}(u, v) = (\cos u, \sin u, \cos v, \sin v), \qquad u, v \in (0, 2\pi).$$

Then the first fundamental form of $\mathbf{T}$ is $du^2 + dv^2$ and we see that $\mathbf{T}$ is locally isometric to the plane. $\mathbf{T}$ is certainly not globally isometric to $\mathbf{R}^2$ since $\mathbf{T}$ is homeomorphic to a torus which is compact and $\mathbf{R}^2$ is non-compact. (In fact the flat torus is isometric to no surface in $\mathbf{R}^3$.) The area of $\mathbf{T}$ is easily seen using (2.4) to equal $4\pi^2$ but as the vector product is not defined in $\mathbf{R}^4$ then our original definition (2.3) is not applicable. $\square$

So far we have only considered examples where the metric structure of the surface is precisely that induced on the surface by the Euclidean space in which the surface lies. There is no reason why we should limit ourselves to these cases – in fact there are good reasons not to.

From Example 44 the tractoid (with the original tractrix removed) has first fundamental form

$$\mathrm{d}s^2 + e^{-2s}\,\mathrm{d}\theta^2, \qquad s > 0, \theta \in (0, 2\pi),$$

when the tractrix is parametrised by arc-length s. The map f from the tractoid to $(0, 2\pi) \times (1, \infty)$ which sends the point on the tractoid with co-ordinates (s, θ) to (θ, e^s) is a diffeomorphism but is not an isometry. We could however ask:

Example 50 *Q: In terms of the co-ordinates x and y find the first fundamental form on $(0, 2\pi) \times (1, \infty)$ for which f is an isometry.*

A: The co-ordinates x and y are related to s and θ by

$$x = \theta, \quad \text{and} \quad y = e^s.$$

For f to be an isometry we need to endow $(0, 2\pi) \times (1, \infty)$ with the first fundamental form

$$\mathrm{d}s^2 + e^{-2s}\,\mathrm{d}\theta^2 = \mathrm{d}(\log y)^2 + \frac{1}{y^2}\mathrm{d}x^2 = \frac{\mathrm{d}x^2 + \mathrm{d}y^2}{y^2}. \quad \square$$

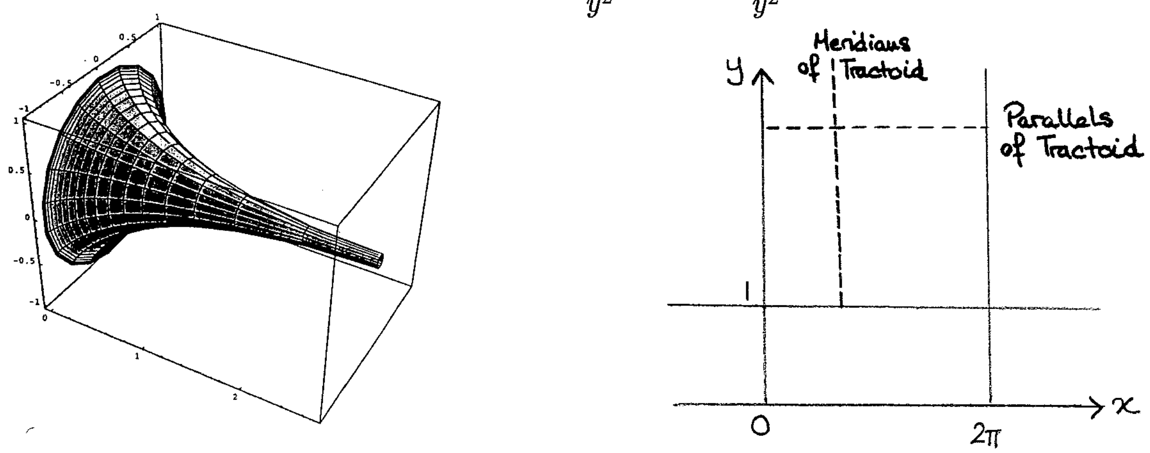


Figure 12 – The Tractoid and Hyperbolic Plane

What we have shown above is that the tractoid (without a meridian) is isometric to part of $\mathbf{H}$, the hyperbolic plane. $\mathbf{H}$ is the surface created by endowing the upper half plane $\{(x, y) : y > 0\}$ with the first fundamental form

$$\frac{\mathrm{d}x^2 + \mathrm{d}y^2}{y^2}. \tag{2.10}$$

$\mathbf{H}$ is of interest because it was the first model for a *non-Euclidean* geometry.

Whilst the infinite rectangle $(0, 2\pi) \times (1, \infty)$ with the first fundamental form (2.10) is isometric to a surface in $\mathbf{R}^3$, the hyperbolic plane is not. We could isometrically embed $\mathbf{H}$ in a higher dimensional Euclidean space, although the isometry may be a little complicated, but there is no need. From our formulas (2.2),(2.4) we may find the length and area of curves and regions in $\mathbf{H}$ without having to be working in a particular Euclidean space. Indeed we could create a surface by endowing any open subset of $\mathbf{R}^2$ with any first fundamental form $E\mathrm{d}x^2 + 2F\mathrm{d}x\mathrm{d}y + G\mathrm{d}y^2$ provided that E, F, G are smooth functions and

$$E > 0, \quad G > 0, \quad EG - F^2 > 0.$$

Conversely any parametrised surface which is diffeomorphic to an open subset of $\mathbf{R}^2$ would be isometric to one of these surfaces.

Example 51 *Q: Find the length of the curve* $\gamma(t) = (0, t)$ *for* $1 \leq t \leq 2$ *in* $\mathbf{H}$.

A: We have $E = G = y^{-2}$ and $F = 0$. Substituting these into (2.2) we find

$$\mathcal{L}(\gamma) = \int_1^2 \sqrt{\frac{1}{t^2}} \mathrm{d}t = [\log t]_1^2 = \log 2. \qquad \square$$

Exercise 52 Show that the surfaces created by endowing $(0, \alpha) \times (0, \infty)$ with the first fundamental form (2.10) are isometric for any $\alpha > 0$.

2.3 Curvature and the Weingarten Map

Let X be a surface in $\mathbf{R^3}$ parametrised by a chart $\mathbf{r} : U \to X$ and let

$$\mathbf{n} = \frac{\mathbf{r}_u \wedge \mathbf{r}_v}{|\mathbf{r}_u \wedge \mathbf{r}_v|}$$

denote a choice of unit normal. We have already studied the curvature of curves in $\mathbf{R}^3$. When $\gamma(s)$ is a curve parametrised by arc-length then the curvature $\kappa(s)$ of γ at the point $\gamma(s)$ is simply the magnitude of $\ddot{\gamma}(s)$.

When looking at a curve in the surface X the vector $\ddot{\gamma}(s)$ has two natural components, a tangential component and a normal component. As $\dot{\gamma}(s)$ is a unit vector for all s we may decompose $\ddot{\gamma}(s)$ in the form:

$$\ddot{\gamma} = k_n \mathbf{n} + k_g (\mathbf{n} \wedge \dot{\gamma}). \tag{2.11}$$

DEFINITION 53 *(a)* $k_n(s)$ *is known as the* normal curvature *of* γ *at* $\gamma(s)$.

(b) $k_g(s)$ *is known as the* geodesic curvature *of* γ *at* $\gamma(s)$.

(c) The tangent vectors of those curves passing through a point $p \in X$ *which have the maximal and minimal normal curvatures are called* principal directions *and their normal curvatures are known as* principal curvatures. *A curve whose tangent vectors are all principal directions is called a* line of curvature.

(d) (*) A curve in X *whose normal curvature is everywhere zero is called an* asymptotic curve.

(e) A curve in X *whose geodesic curvature is everywhere zero is called a* geodesic.

We shall consider for the moment the normal curvature of curves and we shall use this to define a second quadratic form on the tangent plane at a point of X. As we shall see later in §2.4 and Exercise 120 that the geodesics of a surface and the geodesic curvature of a curve at a point depends only on the first fundamental form of the surface and the direction of the curve.

The normal curvature k_n of γ equals $\ddot{\gamma} \cdot \mathbf{n}$. By the chain rule we know

$$\dot{\gamma} = \dot{u}\mathbf{r}_u + \dot{v}\mathbf{r}_v,$$

and applying the chain rule again we find

$$\ddot{\gamma} = \ddot{u}\mathbf{r}_u + \ddot{v}\mathbf{r}_v + \dot{u}^2\mathbf{r}_{uu} + 2\dot{u}\dot{v}\mathbf{r}_{uv} + \dot{v}^2\mathbf{r}_{vv}.$$

Hence the normal curvature $k_n = \ddot{\gamma} \cdot \mathbf{n}$ equals

$$k_n = L\dot{u}^2 + 2M\dot{u}\dot{v} + N\dot{v}^2,$$

where

$$\begin{aligned} L &= \mathbf{r}_{uu} \cdot \mathbf{n} = -\mathbf{r}_u \cdot \mathbf{n}_u, \\ M &= \mathbf{r}_{uv} \cdot \mathbf{n} = -\mathbf{r}_u \cdot \mathbf{n}_v = -\mathbf{r}_v \cdot \mathbf{n}_u, \\ N &= \mathbf{r}_{vv} \cdot \mathbf{n} = -\mathbf{r}_v \cdot \mathbf{n}_v. \end{aligned} \tag{2.12}$$

Note that the alternative expressions for L, M, N come from the differentiating the equations

$$\mathbf{r}_u \cdot \mathbf{n} = 0 = \mathbf{r}_v \cdot \mathbf{n}.$$

DEFINITION 54 *The quadratic form*

$$\alpha \mathbf{r}_u + \beta \mathbf{r}_v \mapsto L\alpha^2 + 2M\alpha\beta + N\beta^2$$

on the tangent plane T_pX is called the second fundamental form *of X.*

Note that many authors including Do Carmo use e, f, g instead of L, M, N for the coefficients of the second fundamental form.

Just as a curve is effectively determined by its curvature and torsion a surface is determined by its first and second fundamental forms. Although the proof of this theorem is far beyond the scope of this course I include a (rough) statement of:

THEOREM 55 *(*): The Fundamental Theorem of the Local Theory of Surfaces.*
Let E, F, G, L, M, N be differentiable functions on an open set $U \subset \mathbf{R}^2$ which satisfy
(i) $E > 0, G > 0, EG - F^2 > 0$,
(ii) certain compatibility equations (Do Carmo p.235).
Then for each $p \in U$ there is an open set $V \subset U$ containing p and a smooth parametrisation $\mathbf{r}(V)$ of a surface in $\mathbf{R}^3$ with $E, 2F, G$ and $L, 2M, N$ as the coefficients of the first and second fundamental forms. Further a second surface $\tilde{\mathbf{r}}(V)$ in $\mathbf{R}^3$ with the same first and second fundamental forms differs from $\mathbf{r}(V)$ only by a rigid motion of $\mathbf{R}^3$.

In order to define the curvature of the surface at a point we need to introduce the Weingarten map. The Weingarten map is the *differential* (see §2.5) of the normal map and consequently is written as $d\mathbf{n}_p$ in some texts. Curvature, for a curve, is a measure of how quickly the tangent is varying. Similarly for a surface we need to investigate how quickly the tangent plane, or equivalently the normal to the surface is varying. Note that as $\mathbf{n} \cdot \mathbf{n} = 1$ then

$$\mathbf{n} \cdot \mathbf{n}_u = 0 = \mathbf{n} \cdot \mathbf{n}_v.$$

Thus $\mathbf{n}_u$ and $\mathbf{n}_v$ are tangents vectors to the surface.

DEFINITION 56 *The* Weingarten map *at the point p is the linear map $W_p : T_pX \to T_pX$ defined by*

$$W_p\mathbf{r}_u = \mathbf{n}_u, \qquad W_p\mathbf{r}_v = \mathbf{n}_v. \tag{2.13}$$

PROPOSITION 57 *The Weingarten map $W_p : T_pX \to T_pX$ is a self-adjoint linear map independent of the choice of parameters u and v. In particular as W_p is self-adjoint it is diagonalisable.*

PROOF: Let $\mathbf{s}(\tilde{u}, \tilde{v})$ be a second chart for X with $\mathbf{s}(\tilde{u}, \tilde{v}) = \mathbf{r}(u, v)$. Then by the chain rule we have

$$\mathbf{s}_{\tilde{u}} = \frac{\partial u}{\partial \tilde{u}}\mathbf{r}_u + \frac{\partial v}{\partial \tilde{u}}\mathbf{r}_v, \qquad \mathbf{s}_{\tilde{v}} = \frac{\partial u}{\partial \tilde{v}}\mathbf{r}_u + \frac{\partial v}{\partial \tilde{v}}\mathbf{r}_v.$$

Hence by the above definition of the Weingarten map (2.13) and the chain rule we have

$$W_p\mathbf{s}_{\tilde{u}} = \frac{\partial u}{\partial \tilde{u}}\mathbf{n}_u + \frac{\partial v}{\partial \tilde{u}}\mathbf{n}_v = \mathbf{n}_{\tilde{u}}, \qquad W_p\mathbf{s}_{\tilde{v}} = \frac{\partial u}{\partial \tilde{v}}\mathbf{n}_u + \frac{\partial v}{\partial \tilde{v}}\mathbf{n}_v = \mathbf{n}_{\tilde{v}}.$$

It is also easy to check that W_p is a self-adjoint linear map – that is

$$(W_p\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (W_p\mathbf{y}) \tag{2.14}$$

for any two tangent vectors $\mathbf{x}, \mathbf{y} \in T_pX$. We note from equation (2.12) that

$$W_p\mathbf{r}_u \cdot \mathbf{r}_v = \mathbf{n}_\mathbf{u} \cdot \mathbf{r}_v = \mathbf{n}_\mathbf{v} \cdot \mathbf{r}_u = W_p\mathbf{r}_v \cdot \mathbf{r}_u.$$

Equation (2.14) then follows for all tangent vectors $\mathbf{x}, \mathbf{y}$ by linearity. $\square$.

In order to work out the eigenvalues and eigenvectors of W_p, let γ be a curve in X with $\gamma(0) = p$. Then

$$\begin{aligned} W_p(\gamma'(0)) \cdot \gamma'(0) &= \mathbf{n}'(\gamma(0)) \cdot \gamma'(0) \\ &= -\mathbf{n} \cdot \gamma''(0) = -k_n. \end{aligned}$$

Thus the eigenvalues of W_p are $-k_1$ and $-k_2$ where k_1 and k_2 are the principal curvatures of X at p and the eigenvectors of W_p are the principal directions (see Definition 53).

We make the following definitions:

DEFINITION 58 *(a) The* Gaussian curvature $K(p)$ *at the point* p *is the product of the principal curvatures or equivalently* $\det W_p$.

(b) (*) The average of the principal curvatures is known as the mean curvature *at* p *or equivalently* $-\frac{1}{2}\mathrm{trace} W_p$.

The tangent vectors $\mathbf{r}_u$ and $\mathbf{r}_v$ form a basis for the tangent plane T_pX and $W_p : T_pX \to T_pX$ is a linear map. So what is the matrix for W_p with respect to this basis?

Let us suppose that the matrix for W_p with respect to the basis $\{\mathbf{r}_u, \mathbf{r}_v\}$ is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Then

$$W_p\mathbf{r}_u = \mathbf{n}_u = A\mathbf{r}_u + C\mathbf{r}_v, \qquad (2.15)$$

$$W_p\mathbf{r}_v = \mathbf{n}_v = B\mathbf{r}_u + D\mathbf{r}_v. \qquad (2.16)$$

Dotting equation (2.15) with $\mathbf{r}_u$ and with $\mathbf{r}_v$ we find

$$-L = AE + CF, \qquad -M = AF + CG.$$

Doing the same for equation (2.16) we obtain

$$-M = BE + DF, \qquad -N = BF + DG.$$

Putting these equations into matrix form gives

$$-\begin{pmatrix} L & M \\ M & N \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

and hence with respect to the basis $\{\mathbf{r}_u, \mathbf{r}_v\}$

$$W_p = \frac{1}{EG - F^2} \begin{pmatrix} -G & F \\ F & -E \end{pmatrix} \begin{pmatrix} L & M \\ M & N \end{pmatrix}.$$

COROLLARY 59 *The Gaussian curvature* $K(p)$ *at* p*, which equals* $\det W_p$*, is given by the formula*

$$K(p) = \frac{LN - M^2}{EG - F^2}.$$

Despite the above expression for K which is in terms of the coefficients of the first *and* second fundamental forms, the Gaussian curvature may be written solely in terms of the coefficients of the first fundamental form and is invariant under isometries. This is a theorem due to Gauss and known as the *Theorema Egregium* (Segal §10).

Gauss originally did not define K by the above formula but rather as the following more intuitive limit. Let U be a small open subset of X about the point p. Then if we let the area of U tend to zero (Segal p.71)

$$|K| = \lim_{\mathrm{Area}(U) \to 0} \frac{\mathrm{Area}(\mathbf{n}(U))}{\mathrm{Area}(U)}.$$

The more 'curved' the surface at a point, the greater the variety in the normal vectors about the point.

We end this section with two worked examples:

Example 60 *Q: The connected smooth parametrised surface* $\mathbf{r}(u, v)$ *has first and second fundamental forms related by the equations*

$$L = hE, \quad M = hF, \quad N = hG \tag{2.17}$$

where $h(u, v)$ *is a smooth function. Show that* h *is constant and that the surface is part of a plane or a sphere.*

A: Let $\mathbf{n}$ be the unit normal to the surface. Recall that

$$E = \mathbf{r}_u \cdot \mathbf{r}_u, \quad F = \mathbf{r}_u \cdot \mathbf{r}_v, \quad G = \mathbf{r}_v \cdot \mathbf{r}_v,$$

and

$$L = -\mathbf{r}_u \cdot \mathbf{n}_u, \quad M = -\mathbf{r}_u \cdot \mathbf{n}_v = -\mathbf{r}_v \cdot \mathbf{n}_u, \quad N = -\mathbf{r}_v \cdot \mathbf{n}_v.$$

Thus the equations (2.17) may be rewritten as

$$(h\mathbf{r}_u + \mathbf{n}_u) \cdot \mathbf{r}_u = 0, \quad (h\mathbf{r}_u + \mathbf{n}_u) \cdot \mathbf{r}_v = 0,$$
$$(h\mathbf{r}_v + \mathbf{n}_v) \cdot \mathbf{r}_u = 0, \quad (h\mathbf{r}_v + \mathbf{n}_v) \cdot \mathbf{r}_v = 0.$$

As $\mathbf{r}_u$ and $\mathbf{r}_v$ form a basis for the tangent plane then

$$h\mathbf{r}_u + \mathbf{n}_u = \mathbf{0} = h\mathbf{r}_v + \mathbf{n}_v. \tag{2.18}$$

If we differentiate the left equation with respect to v and the right with respect to u and subtract the resulting equations then we obtain

$$h_v\mathbf{r}_u - h_u\mathbf{r}_v = \mathbf{0}.$$

As $\mathbf{r}_u$ and $\mathbf{r}_v$ are independent then $h_u = h_v = 0$ and since surface is connected then h is constant.

Putting this back into equations (2.18) gives

$$(h\mathbf{r} + \mathbf{n})_u = \mathbf{0} = (h\mathbf{r} + \mathbf{n})_v$$

and so $h\mathbf{r} + \mathbf{n}$ equals some constant vector $\mathbf{c}$. If $h = 0$ then the normal vector $\mathbf{n}$ is constant and so the surface is part of a plane. If $h \neq 0$ then

$$\mathbf{r} - \frac{\mathbf{c}}{h} = -\frac{\mathbf{n}}{h}$$

and so the surface is part of a sphere, centre $\mathbf{c}/h$ and radius $1/|h|$. □

Example 61 *Q: Consider the parametrised surface (*Enneper's Surface *– see Figure 13),*

$$\mathbf{r}(u, v) = \left(u - \frac{u^3}{3} + uv^2, v - \frac{v^3}{3} + vu^2, u^2 - v^2\right), \qquad u, v \in \mathbf{R}.$$

Find the first and second fundamental forms. Show that the lines of curvature are the co-ordinate curves and that the principal curvatures are $k_1 = 2(1 + u^2 + v^2)^{-2}$ *and* $k_2 = -2(1 + u^2 + v^2)^{-2}$.

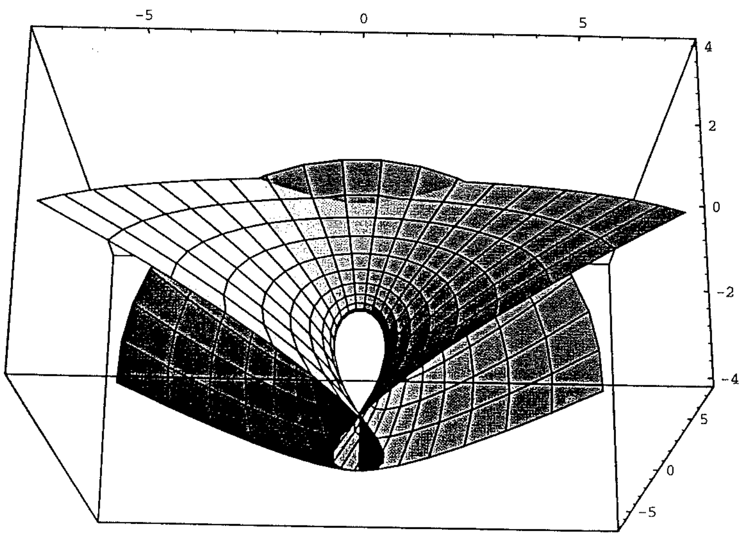


Figure 13 – Enneper's Surface

Remark 62 For an analysis of the self-intersections of Enneper's surface, see Do Carmo p.206.

A: We see

$$\mathbf{r}_u = (1 - u^2 + v^2, 2uv, 2u) \text{ and } \mathbf{r}_v = (2uv, 1 - v^2 + u^2, -2v),$$

giving

$$E = (1 + u^2 + v^2)^2, \quad F = 0, \quad G = (1 + u^2 + v^2)^2.$$

We further find that

$$\mathbf{r}_u \wedge \mathbf{r}_v = (1 + u^2 + v^2) \times (-2u, 2v, 1 - u^2 - v^2)$$

which has magnitude $(1 + u^2 + v^2)^2$. Now

$$\mathbf{r}_{uu} = (-2u, 2v, 2), \quad \mathbf{r}_{uv} = (2v, 2u, 0), \quad \mathbf{r}_{vv} = (2u, -2v, -2),$$

and hence

$$L = 2, \quad M = 0, \quad N = -2.$$

The Weingarten map at the point $\mathbf{r}(u, v)$ equals

$$\frac{1}{(1 + u^2 + v^2)^2} \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix},$$

which has eigenvalues $\pm 2(1 + u^2 + v^2)^{-2}$ and eigenvectors $\mathbf{r}_u, \mathbf{r}_v$ as required.

Notice that the mean curvature of the surface is everywhere zero – such surfaces are known as minimal surfaces. As the name suggests when a minimal surface is perturbed slightly then families of surfaces are produced with greater areas. (See Segal §9.) The catenoid and helicoid (see Example 47) are also examples of minimal surfaces. □

2.4 Geodesics

We gave in the previous section the definition of a geodesic curve (see (2.11)). Namely a geodesic is a curve with zero geodesic curvature or equivalently:

DEFINITION 63 *A curve $\gamma : I \to X$, parametrised by arc-length on a surface X, is a* geodesic *if for all $s \in I$ the vector $\ddot{\gamma}(s)$ is normal to the surface at the point $\gamma(s)$.*

Geodesics are also the curves of shortest length on a surface – at least 'locally'. This means that given a geodesic between two points on a surface, varying the geodesic slightly will produce curves of greater length. For example, given two points on a sphere the great circle containing these two points is a geodesic. If the points are not antipodal then the arcs of the great circle connecting them will be of different lengths. However both arcs are geodesics and locally are the shortest paths between the points.

Unlike asymptotic curves, which are solutions to a differential equation involving the second fundamental form, geodesics are determined by the first fundamental form. Consequently an isometry between two surfaces will map geodesics in the first surface to geodesics in the second.

THEOREM 64 *Let X be a smooth parametrised surface and γ be a smooth curve on X parametrised by arc-length s. Then γ is a geodesic if and only the parameters $(u(s), v(s))$ of $\gamma(s)$ satisfy*

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}(E\dot{u} + F\dot{v}) &= \frac{1}{2}(E_u\dot{u}^2 + 2F_u\dot{u}\dot{v} + G_u\dot{v}^2) \\ \frac{\mathrm{d}}{\mathrm{d}s}(F\dot{u} + G\dot{v}) &= \frac{1}{2}(E_v\dot{u}^2 + 2F_v\dot{u}\dot{v} + G_v\dot{v}^2) \end{aligned} \tag{2.19}$$

for all s, where $E du^2 + 2F du dv + G dv^2$ is the first fundamental form of X.

PROOF: As $\mathbf{r}_u$ and $\mathbf{r}_v$ are independent tangent vectors then $\ddot{\gamma}(s)$ is normal to the surface if and only if $\ddot{\gamma}(s)\cdot\mathbf{r}_u = 0$ and $\ddot{\gamma}(s)\cdot\mathbf{r}_v = 0$. Now

$$\dot{\gamma}(s) = \dot{u}\mathbf{r}_u + \dot{v}\mathbf{r}_v.$$

Thus

$$\begin{aligned}
0 &= \ddot{\gamma}\cdot\mathbf{r}_u = \frac{\mathrm{d}}{\mathrm{d}s}(\dot{\gamma}\cdot\mathbf{r}_u) - \dot{\gamma}\cdot\dot{\mathbf{r}}_u \\
&= \frac{\mathrm{d}}{\mathrm{d}s}(E\dot{u}+F\dot{v}) - (\dot{u}\mathbf{r}_u + \dot{v}\mathbf{r}_v)\cdot(\mathbf{r}_{uu}\dot{u} + \mathbf{r}_{uv}\dot{v}) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}(E\dot{u}+F\dot{v}) - ((\mathbf{r}_{uu}\cdot\mathbf{r}_u)\dot{u}^2 + (\mathbf{r}_{uu}\cdot\mathbf{r}_v + \mathbf{r}_{uv}\cdot\mathbf{r}_u)\dot{u}\dot{v} + (\mathbf{r}_{uv}\cdot\mathbf{r}_v)\dot{v}^2) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}(E\dot{u}+F\dot{v}) - \frac{1}{2}(E_u\dot{u}^2 + 2F_u\dot{u}\dot{v} + G_u\dot{v}^2),
\end{aligned}$$

as required. The second geodesic equation follows similarly. □

Given two points on a surface there need not be a geodesic connecting the two points. For example in $\mathbf{R}^2$ the geodesics are line segments. So in the punctured plane $\mathbf{R}^2 - \{\mathbf{0}\}$ there is no geodesic connecting $(1,0)$ and $(-1,0)$. Also if a geodesic exists between two points it need not be unique (see the examples of the sphere and cylinder below.) However geodesics always exist locally (Do Carmo p.255):

THEOREM 65 *Given a point $p \in X$ and a non-zero vector $\mathbf{v} \in T_pX$ then there exists $\epsilon > 0$ and a unique geodesic $\gamma : (-\epsilon, \epsilon) \to X$ parametrised by arc-length such that $\gamma(0) = p$ and $\gamma'(0) = \mathbf{v}$.*

Example 66 *Q: Determine all the geodesics of a sphere. By solving the geodesic equations above find all the geodesics on the cylinder $x^2 + y^2 = a^2$ in $\mathbf{R}^3$, $(a \neq 0)$.*

A: Without loss of generality assume that the sphere has centre the origin and radius a. Certainly any great circle γ is a geodesic as $a^2\ddot{\gamma}(s) = -\gamma(s) = -\mathbf{n}(\gamma(s))$. We now know from Theorem 65 that these are all the geodesics. Alternatively we can argue as follows. Let γ be a geodesic on the sphere so that

$$0 = \ddot{\gamma}(s)\wedge\mathbf{n}(\gamma(s)) = \ddot{\gamma}(s)\wedge\gamma(s) = \frac{\mathrm{d}}{\mathrm{d}s}(\dot{\gamma}(s)\wedge\gamma(s)).$$

Hence $\dot{\gamma}(s)\wedge\gamma(s) = \mathbf{c}$ for some constant vector $\mathbf{c}$. Then $\gamma(s)\cdot\mathbf{c} = 0$ and so γ lies in a plane through the origin – that is γ is part of a great circle.

We may parametrise the cylinder by setting

$$\mathbf{r}(z,\theta) = (a\cos\theta, a\sin\theta, z), \qquad z \in \mathbf{R}, \theta \in (0, 2\pi).$$

The first fundamental form is then $\mathrm{d}z^2 + a^2\mathrm{d}\theta^2$ and hence the geodesic equations are $\ddot{z} = 0 = \ddot{\theta}$. So

$$z(s) = as + b, \qquad \theta(s) = cs + d$$

for some a, b, c, d where a and c are not both zero. The three cases are then:
(i) $a = 0$: z is constant, and the geodesic is part of a circle,
(ii) $c = 0$: θ is constant, and the geodesic is part of a meridian,
(iii) $a \neq 0, c \neq 0$: the geodesic is part of a helix.
Note that there are infinitely many geodesics between two points on the cylinder which lie on different latitudes. □

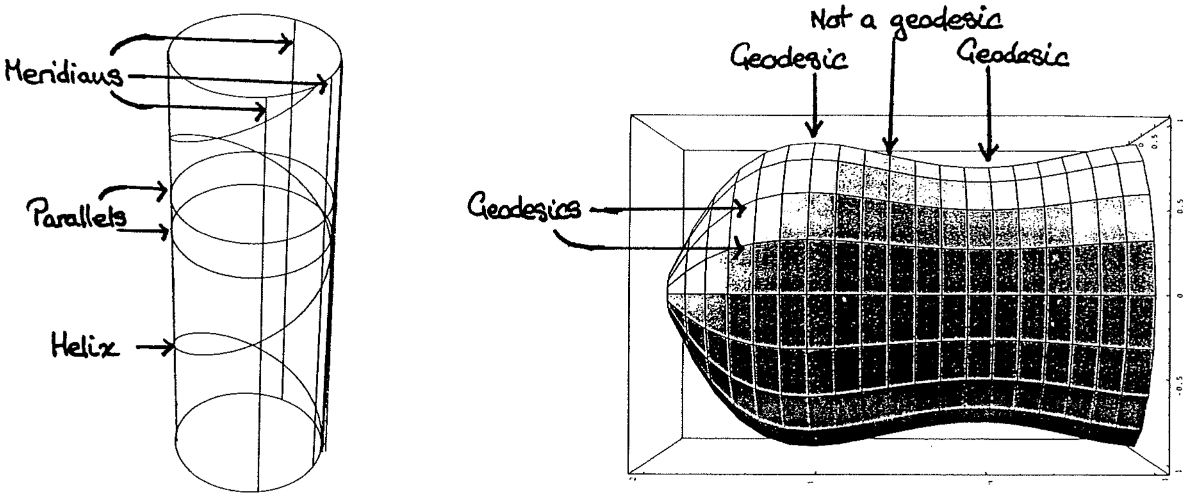


Figure 14 – Geodesics on The Cylinder and Surfaces of Revolution

Example 67 *Q:*
(i) Prove that a meridian on a surface of revolution is a geodesic.
(ii) When is a parallel of latitude a geodesic on such a surface?
(iii) Prove that along a geodesic γ on a surface of revolution the product $f \sin\psi$ is constant, where f is the distance of $\gamma(s)$ from the axis of the surface, and ψ is the angle between $\dot\gamma(s)$ and the meridian through $\gamma(s)$. This is called Clairaut's relation.
(iv) Prove that on the ellipsoid

$$\frac{x^2}{a^2}+\frac{y^2}{a^2}+\frac{z^2}{b^2}=1 \qquad (2.20)$$

every geodesic which is not a meridian remains always between two parallels of latitude.

A: Suppose that the surface of revolution is generated by rotating the curve $y = f(x)$ about the x-axis and parametrise it as

$$\mathbf{r}(x,\theta) = (x, f(x)\cos\theta, f(x)\sin\theta) \qquad x \in \mathbf{R}, \theta \in (-\pi,\pi).$$

The first fundamental form equals

$$(1+(f')^2)\mathrm{d}x^2 + f^2\mathrm{d}\theta^2$$

and the geodesic equations are

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}((1+(f')^2)\dot{x}) &= f'(f''\dot{x}^2 + f\dot\theta^2),\\ \frac{\mathrm{d}}{\mathrm{d}s}(f^2\dot\theta) &= 0.\end{aligned}$$

(i) Along a meridian $\dot\theta = 0$ and $\dot{x} = (1+(f')^2)^{-1/2}$. The second equation is then trivially true and substituting into the first equation we find

$$\frac{\mathrm{d}}{\mathrm{d}s}((1+(f')^2)\dot{x}) = \dot{x}\frac{\mathrm{d}}{\mathrm{d}x}(1+(f')^2)^{1/2} = \frac{f'f''}{1+(f')^2} = f'f''\dot{x}^2$$

as required.

(ii) A parallel is given by the equation $\dot{x} = 0$. Thus the two geodesic equations now read as $ff'\dot\theta^2 = 0$ and $f^2\ddot\theta = 0$. As $f > 0$ and $\dot\theta \neq 0$ then the equations hold if and only if $f' = 0$.

(iii) From the second geodesic equation we can see that $f^2\dot\theta$ is constant along a geodesic. Without any loss of generality suppose that $\theta = 0$ at the point $\gamma(s)$. So the meridian through $\gamma(s)$ is given by $(x, f(x), 0)$ and so the tangent vector to this meridian equals

$$\mathbf{t} = \left(\frac{1}{\sqrt{1+(f')^2}}, \frac{f'}{\sqrt{1+(f')^2}}, 0\right).$$

Further

$$\dot{\gamma}(s) = \dot{x}\mathbf{r}_x + \dot{\theta}\mathbf{r}_\theta = (\dot{x}, f'\dot{x}, f\dot{\theta}).$$

Thus

$$\sin\psi = |\dot{\gamma}(s) \wedge \mathbf{t}| = \frac{1}{\sqrt{1+(f')^2}}\left|\left(-ff'\dot{\theta}, f\dot{\theta}, 0\right)\right| = f\dot{\theta}.$$

Hence $f \sin\psi$ is constant as required.

(iv) Let γ be a geodesic on the ellipsoid (2.20), which is a surface of revolution as two of the ellipsoid's axes are equal, and let h equal the constant value of $f\sin\psi$ on γ. If γ passes through either pole then $h = 0$; at any point of γ other than a pole $\psi = 0$ and hence γ is a meridian. If $h \neq 0$ then $f \geq |f\sin\psi| = |h| > 0$. Hence γ is bounded between two parallels. □

We end with a theorem proving an earlier comment on geodesics namely that they are locally curves of least length. That is, however a geodesic between two points is perturbed we produce families of curves of greater length.

THEOREM 68 *Let $\gamma : [a,b] \to X$ be a smooth geodesic in X. Let γ_s $(s \in (-\epsilon, \epsilon))$ be a family of smooth curves*

$$\gamma_s : [a,b] \to X \quad s \in (-\epsilon, \epsilon)$$

with $\gamma_0 = \gamma$ and $\gamma_s(a) = \gamma(a), \gamma_s(b) = \gamma(b)$ for all $s \in (-\epsilon, \epsilon)$ and let $\mathcal{L}(s) = \mathcal{L}(\gamma_s)$. Then $\mathcal{L}'(0) = 0$.

PROOF: Let $R(s,t) = E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2$ where $\gamma_s(t) = \mathbf{r}(u(s,t), v(s,t))$ and the dot denotes differentiation with respect to t. Then

$$\mathcal{L}(s) = \int_a^b \sqrt{R}\,\mathrm{d}t$$

giving

$$\mathcal{L}'(0) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \int_a^b \sqrt{R}\,\mathrm{d}t = \int_a^b \left.\frac{\partial\sqrt{R}}{\partial s}\right|_{s=0} \mathrm{d}t = \frac{1}{2}\int_a^b \frac{1}{\sqrt{R}} \left.\frac{\partial R}{\partial s}\right|_{s=0} \mathrm{d}t.$$

Now

$$\begin{aligned} \frac{\partial R}{\partial s} &= \{E_u\dot{u}^2 + 2F_u\dot{u}\dot{v} + G_u\dot{v}^2\}\frac{\partial u}{\partial s} \\ &+ \{E_v\dot{u}^2 + 2F_v\dot{u}\dot{v} + G_v\dot{v}^2\}\frac{\partial v}{\partial s} \\ &+ 2(E\dot{u} + F\dot{v})\frac{\partial \dot{u}}{\partial s} + 2(F\dot{u} + G\dot{v})\frac{\partial \dot{v}}{\partial s}. \end{aligned}$$

As $\gamma = \gamma_0$ is a geodesic then substituting in the geodesic equations (2.19)

$$\begin{aligned} \left.\frac{\partial R}{\partial s}\right|_{s=0} &= 2\left[\frac{\mathrm{d}}{\mathrm{d}t}(E\dot{u} + F\dot{v})\frac{\partial u}{\partial s} + (E\dot{u} + F\dot{v})\frac{\partial \dot{u}}{\partial s}\right. \\ &+ \left.\frac{\mathrm{d}}{\mathrm{d}t}(F\dot{u} + G\dot{v})\frac{\partial v}{\partial s} + (F\dot{u} + G\dot{v})\frac{\partial \dot{v}}{\partial s}\right]_{s=0} \\ &= 2\frac{\mathrm{d}}{\mathrm{d}t}\left\{(E\dot{u} + F\dot{v})\left.\frac{\partial u}{\partial s}\right|_{s=0} + (F\dot{u} + g\dot{v})\left.\frac{\partial v}{\partial s}\right|_{s=0}\right\}. \end{aligned}$$

We may assume without loss of generality that $\gamma = \gamma_0$ is parametrised by arc-length so that $R(0,t) = 1$. Hence

$$\mathcal{L}'(0) = \int_a^b \frac{\mathrm{d}}{\mathrm{d}t}\left\{(E\dot{u} + F\dot{v})\left.\frac{\partial u}{\partial s}\right|_{s=0} + (F\dot{u} + g\dot{v})\left.\frac{\partial v}{\partial s}\right|_{s=0}\right\}\mathrm{d}t$$

which equals

$$\left[(E\dot{u} + F\dot{v})\left.\frac{\partial u}{\partial s}\right|_{s=0} + (F\dot{u} + g\dot{v})\left.\frac{\partial v}{\partial s}\right|_{s=0}\right]_{t=a}^{t=b}.$$

However $u(s,a), u(s,b), v(s,a)$ and $v(s,b)$ are all constant giving $\partial u/\partial s = \partial v/\partial s = 0$ when $t = a$ and $t = b$ and hence $\mathcal{L}'(0) = 0$. □

2.5 Further Topics (*)

2.5.1 Atlases

Everything covered in this closing section is beyond the scope of the a2 course. These topics largely concern moves towards the global geometry of surfaces (we have previously only been concerned with local properties) and abstract surfaces.

Thus far we have only considered surfaces in $\mathbf{R}^3$ which are diffeomorphic to $\mathbf{R}^2$. This excludes surfaces such as the sphere or cylinder; in these cases we were satisfied to parametrise dense open subsets and this was satisfactory while we were solely concerned with their local geometry.

To consider the global geometry of a surface we need to be able to parametrise the entire surface – to do this we simply use more parametrisations, the union of whose images is the entire surface. Such a collection of charts is called an *atlas* for the surface.

DEFINITION 69 *A subspace $X \subset \mathbf{R}^3$ is a* smooth surface *if about every $p \in X$ there is an open neighbourhood $V \subseteq X$ and a chart $\mathbf{r} : U \to V$ from an open subset $U \subseteq \mathbf{R}^2$ such that $\mathbf{r} : U \to V$ is a smooth parametrised surface.*

Throughout these notes we have shown that such properties of the surface as smoothness, curvature, etc. are independent of the choice of parametrisation. Consequently, though a point of the surface X may lie in the image of two charts, we may discuss (for example) the curvature of a surface at a point without reference to a specific chart.

Example 70 *Q: Find an atlas for the unit sphere $x^2 + y^2 + z^2 = 1$.*

A: We will find six charts for the sphere whose images are the six hemispheres which lie to either side of the xy-, yz- and xz-planes. So define:

$$\begin{aligned}
\mathbf{r}_1(u,v) &= (u, v, \sqrt{1-u^2-v^2}) \quad & u^2+v^2<1,\\
\mathbf{r}_2(u,v) &= (u, v, -\sqrt{1-u^2-v^2}) \quad & u^2+v^2<1,\\
\mathbf{r}_3(u,v) &= (u, \sqrt{1-u^2-v^2}, v) \quad & u^2+v^2<1,\\
\mathbf{r}_4(u,v) &= (u, -\sqrt{1-u^2-v^2}, v) \quad & u^2+v^2<1,\\
\mathbf{r}_5(u,v) &= (\sqrt{1-u^2-v^2}, u, v) \quad & u^2+v^2<1,\\
\mathbf{r}_6(u,v) &= (-\sqrt{1-u^2-v^2}, u, v) \quad & u^2+v^2<1.
\end{aligned}$$

We can easily check that each of the maps above is a smooth parametrised surface and so the sphere is indeed a smooth surface. □

Exercise 71 Find atlases for the following surfaces:
(a) The cylinder $x^2 + y^2 = a^2$ where $a > 0$.
(b) The hyperboloid of two sheets $x^2 + y^2 + 1 = z^2$.
(c) The flat torus $\{(x, y, z, t) \in \mathbf{R}^4 : x^2 + y^2 = z^2 + t^2 = 1\}$.

We discussed earlier in §2.1 the notions of smooth maps between parametrised surfaces and of diffeomorphisms. As a smooth surface can locally be parametrised then we may extend our notion of smooth maps between parametrised surfaces more generally to smooth surfaces.

DEFINITION 72 *Let X and Y be smooth surfaces in $\mathbf{R}^3$ with respective atlases*

$$\{\mathbf{r}_\alpha : U_\alpha \to V_\alpha \subseteq X, \alpha \in A\}, \quad \{\tilde{\mathbf{r}}_\beta : \tilde{U}_\beta \to \tilde{V}_\beta \subseteq Y, \beta \in B\}.$$

Then we say that a map $f : X \to Y$ is smooth *if*

$$\tilde{\mathbf{r}}_\beta^{-1} \circ f \circ \mathbf{r}_\alpha$$

is smooth whenever $f(V_\alpha) \cap \tilde{V}_\beta \neq \emptyset$. A map $f : X \to Y$ which is smooth with a smooth inverse is called a diffeomorphism.

We noted earlier that all parametrised surfaces are diffeomorphic to one another. This is certainly no longer the case with smooth surfaces – no pair of the sphere, torus, cylinder, plane are diffeomorphic because no such pair is homeomorphic.

Exercise 73 Construct a diffeomorphism between the sphere $x^2+y^2+z^2=1$ and the ellipsoid $x^2/a^2+y^2/b^2+z^2/c^2=1$ where $a,b,c>0$.

Exercise 74 Let X be a surface of revolution. Show that a rotation of X about its axis is a diffeomorphism of X.

Now that we have defined what is meant by an atlas of a surface and have co-ordinate systems over entire smooth surfaces we can introduce our first example of a global property of surfaces. The plane $\mathbf{R}^2$ has more 'structure' than simply that of a smooth surface – it has a metric structure and also a notion of sense, i.e. we think about the vectors $\{\mathbf{i},\mathbf{j}\}$ (in that order) as being a right-handed basis and we have notions of clockwise/anti-clockwise. We can define for a particular co-ordinate system on a parametrised surface these notions as well. In order to be able to do this for a general smooth surface we need to have consistent notions of sense between the different charts of the atlas. Consequently we are interested in the *transition maps* of an atlas.

DEFINITION 75 *Let X be a smooth surface and let*

$$\{\mathbf{r}_\alpha : U_\alpha \to V_\alpha \subseteq X, \alpha \in A\},$$

be an atlas for this surface. Then the transition maps *(see Figure 15) of the atlas are those maps*

$$(\mathbf{r}_\beta)^{-1} \circ \mathbf{r}_\alpha : (\mathbf{r}_\alpha)^{-1}(V_\alpha \cap V_\beta) \to (\mathbf{r}_\beta)^{-1}(V_\alpha \cap V_\beta)$$

which are defined whenever $V_\alpha \cap V_\beta \neq \emptyset$.

The transition maps of a smooth surface in $\mathbf{R}^3$ are diffeomorphisms (Do Carmo p.70) and this is because the smooth structure of the surface is consistent between charts. If we want to have a consistent notion of 'sense' or *orientation* then we will require that the transition maps are *orientation preserving.*

Let $f : \mathbf{R}^2 \to \mathbf{R}^2$ be a diffeomorphism. Then f maps the x-axis and y-axis to smooth curves in $\mathbf{R}^2$. The tangent vectors to the x- and y-axes at $(0,0)$ are $\mathbf{i},\mathbf{j}$ and the tangent vectors to their images under f (which meet at $f(0,0)$) are

$$\left(\frac{\partial f_1}{\partial x}(0,0), \frac{\partial f_2}{\partial x}(0,0)\right), \qquad \left(\frac{\partial f_1}{\partial y}(0,0), \frac{\partial f_2}{\partial y}(0,0)\right)$$

where $f=(f_1,f_2)$. These two vectors will have the same sense as $\mathbf{i},\mathbf{j}$ if

$$\det \begin{pmatrix} \frac{\partial f_1}{\partial x}(0,0) & \frac{\partial f_1}{\partial y}(0,0) \\ \frac{\partial f_2}{\partial x}(0,0) & \frac{\partial f_2}{\partial y}(0,0) \end{pmatrix} > 0. \tag{2.21}$$

DEFINITION 76 *A diffeomorphism $f : \mathbf{R}^2 \to \mathbf{R}^2$ is said to be* orientation preserving *at $(0,0)$ (or similarly at any other point $p \in \mathbf{R}^2$) if the determinant (2.21) is strictly positive. f is said to be orientation preserving if it is orientation preserving at every point of $\mathbf{R}^2$.*

DEFINITION 77 *A smooth surface $X \subset \mathbf{R}^3$ is* orientable *if there exists an atlas for X with orientation preserving transition maps.*

Not all surfaces are orientable. The projective plane $\mathbf{P}$ described in the next section is not (though we shall not prove this). Clearly every parametrised surface is orientable – as there is only one chart and hence no transition maps in its atlas. Also all compact surfaces in $\mathbf{R}^3$ are orientable. For a surface in $\mathbf{R}^3$ orientability may be rephrased in terms of the normal map. A smooth surface $X \subset \mathbf{R}^3$ is orientable if there is a smooth map $\mathbf{n} : X \to S^2$ such that $\mathbf{n}(p)$ is normal to X at p for each $p \in X$. That is, X is orientable if we can smoothly define a choice of unit normal over the entire surface X.

Exercise 78 Show that the transition maps for the atlas of the sphere in Example 70 are orientation preserving.

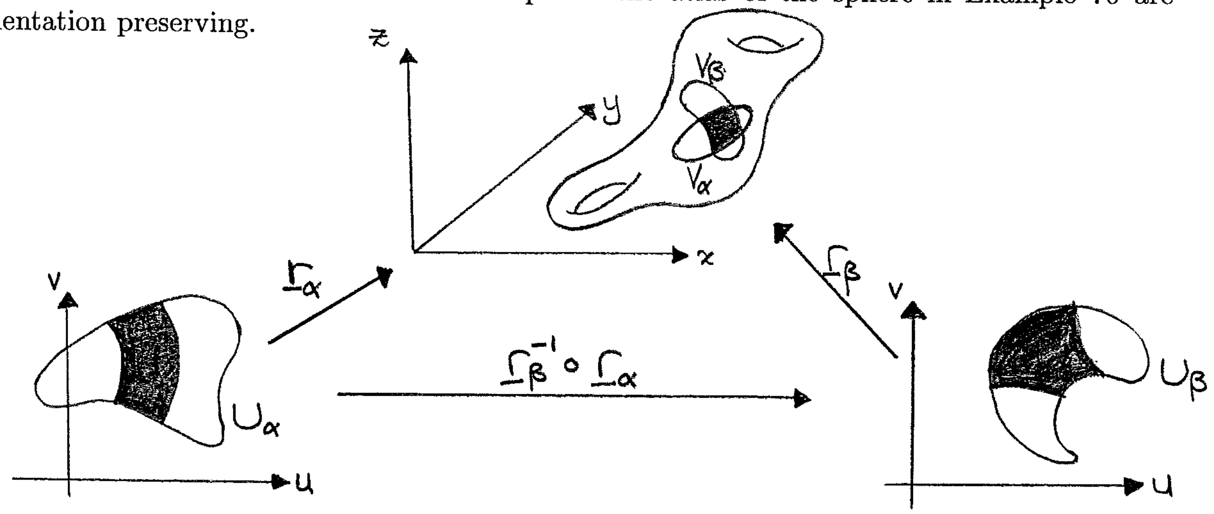
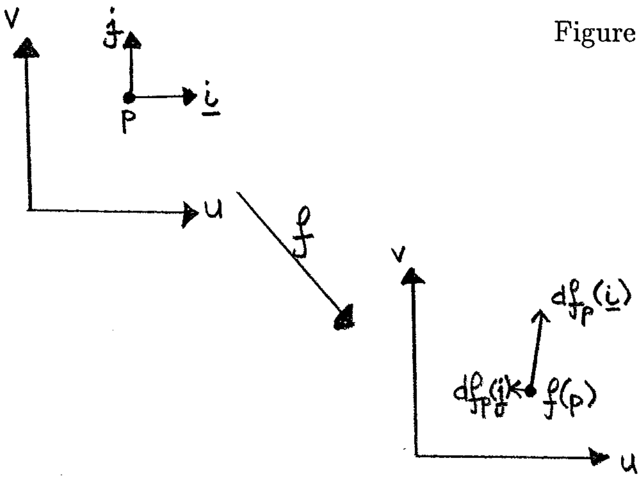


Figure 15



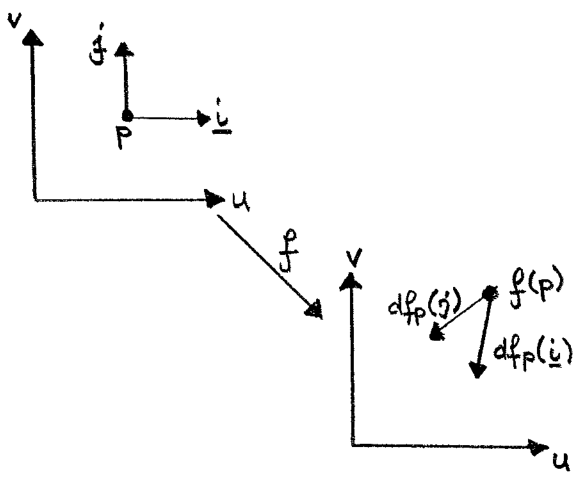


Figure 16 – An Orientation Preserving Map Figure 17 – An Orientation Reversing Map

2.5.2 Abstract and Geometric Surfaces

In the examples we have considered so far we have generally limited ourselves to examples of surfaces lying in Euclidean space. For such surfaces their smooth structure is endowed straight from the ambient Euclidean space. In the example of the hyperbolic plane (see the end of §2.2) we introduced a metric structure to the plane which differed from that of the ambient space. Similarly we can introduce smooth structures that differ from that of the ambient space or even introduce them where no notion of smoothness previously existed.

Example 79 *Q: The* real projective plane $\mathbf{P} = S^2/\{\pm 1\}$ *is the space formed by identifying antipodal points on the sphere. Find an atlas for* $\mathbf{P}$.

A: Each equivalence class of points in $S^2/\{\pm 1\}$ has a representative in one (or more) of the images of the charts $\mathbf{r}_1, \mathbf{r}_3, \mathbf{r}_5$ (see Example 70). Let

$$\pi : S^2 \to \mathbf{P} : x \mapsto \{\pm x\}$$

denote the natural map and then the maps

$$\mathbf{s}_1 = \pi \circ \mathbf{r}_1, \quad \mathbf{s}_3 = \pi \circ \mathbf{r}_3, \quad \mathbf{s}_5 = \pi \circ \mathbf{r}_5, \tag{2.22}$$

form an atlas for $\mathbf{P}$. □

Does the example above really make any sense? We have found co-ordinate systems which cover the whole of $\mathbf{P}$ but one of the requirements for (2.22) to form an atlas is that these maps be smooth. How can we make any sense of that when $\mathbf{P}$ only exists as a topological space, rather than as a subspace of some Euclidean space?

The way around this problem is to *define* the maps (2.22) to be smooth. Given a topological space X and a homeomorphism $\mathbf{r} : U \to X$ from an open set $U \subseteq \mathbf{R}^2$ we may give X a smooth structure by requiring $\mathbf{r}$ to be a diffeomorphism so that $\mathbf{r}$ transfers both the topological structure *and* smooth structure of U to X.

There is a problem of consistency with this approach. If we require $\mathbf{r} : U \to X$ to be a diffeomorphism then we may define

$$f : X \to \mathbf{R}$$

to be a smooth function if $f \circ \mathbf{r} : U \to \mathbf{R}$ is smooth. However in our atlas some points belong in the images of two charts – might not a function be smooth at a point with respect to one chart, but not with respect to another? This situation will not arise if the transition maps of the atlas are smooth. The transition maps for our atlas of $\mathbf{P}$ are the maps

$$(\mathbf{s}_1)^{-1} \circ \mathbf{s}_3, \quad (\mathbf{s}_3)^{-1} \circ \mathbf{s}_5, \quad (\mathbf{s}_5)^{-1} \circ \mathbf{s}_1,$$

and their inverses. If the transition maps of an atlas are smooth then there is agreement as to what functions are smooth at points which lie in the images of two or more charts. Previously when we were considering smooth surfaces in a Euclidean space then the smoothness of the transition maps was automatic because of the ambient space. As we have not actually embedded $\mathbf{P}$ in a Euclidean space we will need to check that they are smooth. The transition maps above are quite complicated, for example,

$$(\mathbf{s}_1)^{-1} \circ \mathbf{s}_3(u,v) = \begin{cases} (u, \sqrt{1-u^2-v^2}), & v > 0, \\ (-u, -\sqrt{1-u^2-v^2}), & v < 0, \end{cases}$$

but nonetheless these maps are smooth. A *smooth abstract surface* or *smooth 2-manifold* is a space with an atlas of charts whose transition maps are all smooth – the smooth structure of the space is then transferred consistently from $\mathbf{R}^2$ by these charts.

DEFINITION 80 *A* smooth abstract surface *is a Hausdorff topological space X together with homeomorphisms $\phi_\alpha : U_\alpha \to V_\alpha$ between open sets $U_\alpha \subseteq \mathbf{R}^2$ and open sets $V_\alpha \subseteq X$ such that*

(a) $\bigcup_\alpha V_\alpha = X$,

(b) when $V_\alpha \cap V_\beta \neq \emptyset$ then

$$(\phi_\alpha)^{-1} \circ \phi_\beta : (\phi_\beta)^{-1}(V_\alpha \cap V_\beta) \to (\phi_\alpha)^{-1}(V_\alpha \cap V_\beta)$$

is smooth.

So $\mathbf{P}$ above is our first example of a smooth abstract surface.

We know already that we can go further and endow a metric structure on an abstract surface. We introduced in §2.2 the hyperbolic plane $\mathbf{H}$ as the upper half plane $\{(x,y) : y > 0\}$ endowed with the first fundamental form

$$\frac{\mathrm{d}x^2 + \mathrm{d}y^2}{y^2} \tag{2.23}$$

So the smooth structure of $\mathbf{H}$ is no different from that of $\mathbf{R}^2$ but we have introduced a different metric structure to the upper half plane. As the atlas for $\mathbf{H}$ consists of only one chart we need not for the moment worry about the consistency of a metric structure on an abstract surface. Having defined the first fundamental form then we are free to ask questions concerning the intrinsic geometry of $\mathbf{H}$.

Example 81 *Q: What are the geodesics in* $\mathbf{H}$*?*

A1: If we substitute $E = G = y^{-2}$ and $F = 0$ into the geodesic equations (2.19) then we find

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\dot{x}}{y^2}\right) = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\dot{y}}{y^2}\right) = \frac{-(\dot{x}^2 + \dot{y}^2)}{y^3}.$$

The first equation yields $\dot{x} = cy^2$ for some constant c. Hence the half-lines $x = \text{const.}$ are geodesics corresponding to $c = 0$. Assume that $c \neq 0$. The second equation may be rewritten as

$$\frac{\ddot{y}y - \dot{y}^2}{y^2} = \frac{-\dot{x}^2}{y^2},$$

or equivalently

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\dot{y}}{y}\right) = -c\dot{x}.$$

Integrating we find that $\dot{y} = (b - cx)y$ for some constant b. Now

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\dot{y}}{\dot{x}} = \frac{b - cx}{cy},$$

and solving this differential equation gives

$$\frac{1}{2}c(x^2 + y^2) - bx = a,$$

for some constant a, which is the equation of a semicircle in $\mathbf{H}$ which cuts the x-axis orthogonally.

A2: Alternatively we could consider what the isometries of $\mathbf{H}$ might be and use the fact that geodesics are mapped to other geodesics by isometries. For ease of notation we now introduce a complex variable $z = x + iy$ so that the first fundamental form (2.23) on $\mathbf{H}$ is now given by

$$\frac{-4|\mathrm{d}z|^2}{(z - \bar{z})^2}.$$

Then I claim the map

$$w : z \mapsto \frac{az + b}{cz + d},$$

where a, b, c, d are real numbers satisfying $ad - bc = 1$, is an isometry of $\mathbf{H}$. From standard theorems concerning Möbius transformations we can see that w maps the upper half plane onto the upper half plane. To check w is an isometry we need to prove that $\mathbf{H}$ when parametrised by w and z has the same first fundamental form. Firstly note

$$\mathrm{d}w = \frac{\mathrm{d}z}{(cz + d)^2}.$$

So

$$\frac{-4|\mathrm{d}w|^2}{(w - \bar{w})^2} = \frac{\dfrac{-4|\mathrm{d}z|^2}{|cz + d|^4}}{\left(\dfrac{az + b}{cz + d} - \dfrac{a\bar{z} + b}{c\bar{z} + d}\right)^2} = \frac{-4|\mathrm{d}z|^2}{((az + b)(c\bar{z} + d) - (cz + d)(a\bar{z} + b))^2}.$$

The denominator in the final expression above factorises as $(ad - bc)^2(z - \bar{z})^2$ showing that

$$\frac{-4|\mathrm{d}w|^2}{(w - \bar{w})^2} = \frac{-4|\mathrm{d}z|^2}{(z - \bar{z})^2}$$

and consequently w is an isometry.

Note now that $x = 0, y = e^{-s}$ is a solution to the geodesic equations for $\mathbf{H}$ and so the positive imaginary axis is a geodesic. Using a Möbius transformation such as w we may map the positive imaginary axis to any other half line or semicircle orthogonal to the real axis, and so these too are examples of geodesics. From Theorem 65 we know that these are all the geodesics of $\mathbf{H}$. □

The hyperbolic plane $\mathbf{H}$ is of interest because it is an example of a *non-Euclidean geometry.* A Euclidean geometry is one that satisfies certain axioms including the *axiom of parallels* which states that:

- given a line l and a point p not on l then there is a unique line through p (known as a parallel) which does not meet l.

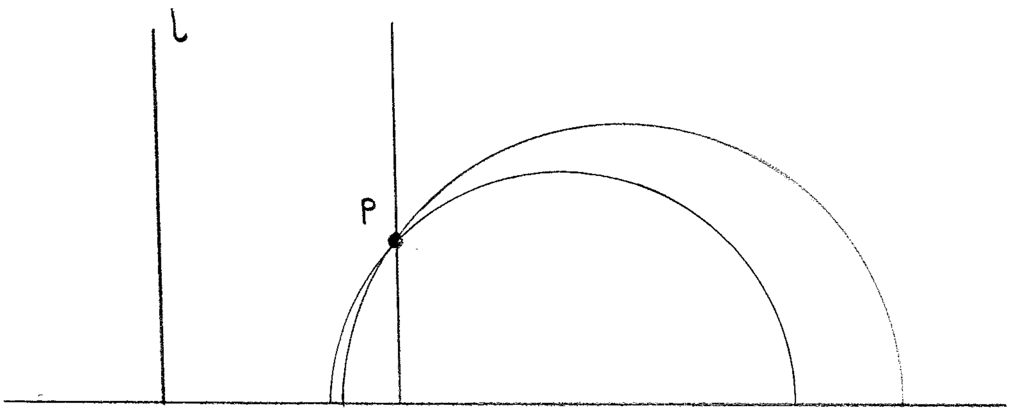


Figure 18 – The Geodesics of The Hyperbolic Plane

If we read 'geodesic' for 'line' in the above then we see that given a line l in $\mathbf{H}$ and a point p not on the line then there are infinitely many lines through p not meeting l.

For centuries mathematicians had been trying to deduce the axiom of parallels from Euclid's other axioms. It was a revolution in geometry when a model for the hyperbolic plane was found (due to Poincaré) which showed that a non-Euclidean geometry existed just as consistent as Euclidean geometry.

In our example of the hyperbolic plane there is only one chart and so the question of the consistency of a metric structure has not been raised. When we were concerned with consistent notions of smoothness, orientation, etc. then we required the transition maps to be diffeomorphisms, orientation preserving etc. Now that we are interested in a consistent metric structure we require these transition maps to be isometries. Hence an abstract geometric surface is defined as:

DEFINITION 82 *A* smooth geometric surface *or* smooth Riemannian 2-manifold *is a Hausdorff topological space X together with*

(i) homeomorphisms $\phi_\alpha : U_\alpha \to V_\alpha$ between open sets $U_\alpha \subseteq \mathbf{R}^2$ and open sets $V_\alpha \subseteq X$,

(ii) first fundamental forms $E_\alpha \mathrm{d}x^2 + 2F_\alpha \mathrm{d}x\mathrm{d}y + G_\alpha \mathrm{d}y^2$ on U_α where $E_\alpha, F_\alpha, G_\alpha$ are smooth functions satisfying

$$E_\alpha > 0, \quad G_\alpha > 0, \quad E_\alpha G_\alpha - (F_\alpha)^2 > 0,$$

such that

(a) $\bigcup_\alpha V_\alpha = X$,

(b) when $V_\alpha \cap V_\beta \neq \emptyset$ then

$$(\phi_\alpha)^{-1} \circ \phi_\beta : (\phi_\beta)^{-1}(V_\alpha \cap V_\beta) \to (\phi_\alpha)^{-1}(V_\alpha \cap V_\beta)$$

is an isometry.

Example 83 *Q: Let D denote the unit disc $\{(u,v) : u^2+v^2 < 1\}$. Find first fundamental forms on $\mathrm{s}_1(D), \mathrm{s}_3(D), \mathrm{s}_5(D)$ such that the natural map $\pi : S^2 \to \mathbf{P}$ is a local isometry.*

A: The chart $\mathbf{r}_1 : D \to S^2$ is defined by

$$\mathbf{r}_1(u,v) = (u, v, \sqrt{1-u^2-v^2}).$$

So the first fundamental form of $\mathbf{r}_1(D)$ is

$$\frac{(1-v^2)\mathrm{d}u^2 + 2uv\mathrm{d}u\mathrm{d}v + (1-u^2)\mathrm{d}v^2}{1-u^2-v^2}.$$

We need to endow $\mathrm{s}_1(D)$ with the above first fundamental form for $\pi : S^2 \to \mathbf{P}$ to be a local isometry. By symmetry we must similarly endow $\mathrm{s}_3(D)$ and $\mathrm{s}_5(D)$ with formally the same first fundamental form. □

Exercise 84 Prove that the transition maps of the above atlas are isometries.

The projective plane endowed with the above first fundamental form is known as the *elliptic plane* and is a further example of a non-Euclidean geometry. As π is a local isometry then the geodesics of the elliptic plane are the images of the great circles. So any two lines of the elliptic plane meet in a unique point – this time no parallels exist.

2.5.3 Differentiability in $\mathbf{R}^n$

We end with a brief discussion of what is meant by the differential of a smooth map between two surfaces.

DEFINITION 85 *Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be a smooth map, (i.e. all partial derivatives of f of all orders exist everywhere.) Let $p, v \in \mathbf{R}^n$ and let $\gamma : (-\epsilon, \epsilon) \to \mathbf{R}^n$ be a smooth curve in $\mathbf{R}^n$ such that*

$$\gamma(0) = p \text{ and } \gamma'(0) = v.$$

Then $f \circ \gamma$ is a smooth curve in $\mathbf{R}^m$. The differential *of f at p is the linear map $df_p : \mathbf{R}^n \to \mathbf{R}^m$ defined by*

$$df_p(v) = df_p(\gamma'(0)) = (f \circ \gamma)'(0).$$

PROPOSITION 86 *$df_p(v)$ is independent of the choice of curve γ.*

PROOF: For ease of notation we shall consider the case when $m = n = 2$. Write $f = \binom{f_1}{f_2}$ and $\gamma = \binom{\gamma_1}{\gamma_2}$. Then

$$\begin{aligned}
(f \circ \gamma)'(0) &= \begin{pmatrix} (f_1 \circ \gamma)'(0) \\ (f_2 \circ \gamma)'(0) \end{pmatrix} \\
&= \begin{pmatrix} \dfrac{\partial f_1}{\partial x}\gamma_1'(0) + \dfrac{\partial f_1}{\partial y}\gamma_2'(0) \\ \dfrac{\partial f_2}{\partial x}\gamma_1'(0) + \dfrac{\partial f_2}{\partial y}\gamma_2'(0) \end{pmatrix} \\
&= \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.
\end{aligned}$$

As the partial derivatives in the above matrix depend only on the function f and the point p then df_p (which we see has the Jacobian as its matrix) is independent of the choice of γ. $\square$

The definition of the differential of a map $f : \mathbf{R}^2 \to \mathbf{R}^2$ extends to maps between parametrised surfaces in an obvious way. Let X and Y be smooth parametrised surfaces and let $p \in X$. Let $f : X \to Y$ be a smooth map. Then the differential of f at p is the linear map

$$\mathrm{d}f_p : T_pX \to T_{f(p)}(Y)$$

defined as follows. Let $v \in T_pX$ and let $\gamma : (-\epsilon, \epsilon) \to X$ be a smooth curve such that

$$\gamma(0) = p \text{ and } \gamma'(0) = v.$$

Then $f \circ \gamma$ is a smooth curve in Y and as before we define

$$df_p(v) = df_p(\gamma'(0)) = (f \circ \gamma)'(0).$$

We have already met examples of differentials. The Weingarten map (see §2.3) is the differential of the normal map. Also this notion of a differential explains the notation du and dv in the first fundamental forms we have studied. Let $X = \mathbf{r}(U)$ be a smooth parametrised surface. Let

$$u : \mathbf{r}(u, v) \mapsto u \text{ and } v : \mathbf{r}(u, v) \mapsto v$$

denote the co-ordinate maps. For $p = \mathbf{r}(u_0, v_0)$ consider the differentials $\mathrm{d}u_p, \mathrm{d}v_p : T_pX \to \mathbf{R}$. We define two curves along the co-ordinate curves through p. Set

$$\begin{aligned} \gamma(t) &= \mathbf{r}(u_0 + t, v_0), \qquad t \in (-\epsilon, \epsilon), \\ \Gamma(t) &= \mathbf{r}(u_0, v_0 + t), \qquad t \in (-\epsilon, \epsilon). \end{aligned}$$

Note that $\gamma'(0) = \mathbf{r}_u(p)$ and $\Gamma'(0) = \mathbf{r}_v(p)$. So

$$\mathrm{d}u_p(\mathbf{r}_u) = \mathrm{d}u_p(\gamma'(0)) = (u \circ \gamma)'(0) = (t \mapsto u_0 + t)'(0) = 1,$$
$$\mathrm{d}u_p(\mathbf{r}_v) = \mathrm{d}u_p(\Gamma'(0)) = (u \circ \Gamma)'(0) = (t \mapsto u_0)'(0) = 0.$$

Similarly $\mathrm{d}v_p(\mathbf{r}_u) = 0$ and $\mathrm{d}u_p(\mathbf{r}_v) = 1$. So $\mathrm{d}u_p$ and $\mathrm{d}v_p$ are elements of the dual of the tangent plane T_pX and are the dual basis of $\{\mathbf{r}_u(p), \mathbf{r}_v(p)\}$. So $E\mathrm{d}u_p^2 + 2F\mathrm{d}u_p\mathrm{d}v_p + G\mathrm{d}v_p^2$ is the quadratic form on T_pX given by

$$\alpha\mathbf{r}_u + \beta\mathbf{r}_v \mapsto E\alpha^2 + 2F\alpha\beta + G\beta^2.$$

Example 87 *Q: Let V denote the space of $n \times n$ real matrices regarded as identified with $\mathbf{R}^{n^2}$. Show that the determinant function $X \mapsto \det X$ from V to $\mathbf{R}$ is differentiable, and prove that its derivative at the identity matrix I is the map $X \mapsto$ trace X.*
For each $X \in V$ let

$$\exp X = I + \sum_{m=1}^{\infty} \frac{X^m}{m!}.$$

[You may assume that the series converges for all $V \in X$.] Find the derivative of the map $X \mapsto \exp X$ at $X = 0$.

For a fixed $A \in V$ let

$$\phi(t) = \det(\exp(tA)).$$

Obtain and solve a differential equation for $\phi(t)$, and hence show that

$$\det(\exp A) = \exp(\mathit{trace}\ A).$$

[You may use without proof the fact that $\exp(X+Y) = \exp(X)\exp(Y)$ for commuting matrices $X, Y \in V$.]

A: det is clearly a smooth function as it is a polynomial of degree n in n^2 variables. Let $X \in V$ and note that

$$\det(I + hX) = 1 + h\mathrm{trace}\ X + O(h^2).$$

Define $\gamma(h) = I + hX$ for $h \in \mathbf{R}$. Then γ is a curve in V with $\gamma(0) = I$ and $\gamma'(0) = X$. Hence

$$\mathrm{d}(\det)_I(X) = (\det\gamma)'(0) = \lim_{h\to 0} \frac{\det(I + hX) - \det(I)}{h} = \mathrm{trace}\ X.$$

Similarly $\Gamma(h) = X$ for $h \in \mathbf{R}$ is a curve in V with $\gamma(0) = 0$ and $\gamma'(0) = X$. Hence

$$\mathrm{d}(\exp)_0(X) = (\exp\gamma)'(0) = \lim_{h\to 0} \frac{\exp(hX) - \exp(0)}{h} = \lim_{h\to 0} \sum_{m=1}^{\infty} \frac{h^{m-1}X^m}{m!} = X.$$

We define $\phi(t) = \det(\exp(tA))$. Then

$$\begin{aligned} \phi'(t) &= \lim_{h\to 0} \frac{\det(\exp((t+h)A)) - \det(\exp(tA))}{h} \\ &= \det(\exp(tA)) \lim_{h\to 0} \frac{\det(\exp(hA)) - 1}{h} = \phi(t)\phi'(0), \end{aligned}$$

which has solution $\phi(t) = \phi(0)\exp(\phi'(0)t)$.

Now $\phi(0) = I$ and by the chain rule (see Exercise 133)

$$\phi'(0) = \mathrm{d}(\det)_I \circ \mathrm{d}(\exp)_0(A) = \mathrm{trace}\ A.$$

Setting $t = 1$ we obtain

$$\det(\exp A) = \exp(\mathrm{trace}\ A). \qquad \square$$

2.6 Exercises

Parametrised Surfaces

Exercise 88 Find smooth parametrisations for dense open subsets of the following surfaces:
(a) The paraboloid $\{(x,y,z) : z = x^2 + y^2\}$,
(b) The ellipsoid $\{(x,y,z) : x^2/a^2 + y^2/b^2 + z^2/c^2 = 1\}$,
(c) The helicoid $\{(x,y,z) : y \sin x = z \cos x\}$.
For each of your parametrisations $\mathbf{r}(U)$ find the image of the normal map $\mathbf{n}(\mathbf{r}(U))$ in the unit sphere.

Exercise 89 Show that if all the normals to a connected surface pass through a fixed point then the surface is contained in a sphere. Show that if all the normals to a connected surface meet a fixed line then the surface is contained in a surface of revolution.

Exercise 90 Let a, b, c be real non-zero constants. Show that the surfaces given by

$$x^2 + y^2 + z^2 = ax, \qquad x^2 + y^2 + z^2 = by, \qquad x^2 + y^2 + z^2 = cz,$$

all intersect orthogonally.

Exercise 91 The equation $x^2 + y^2 = 1$ defines a surface X in $\mathbf{C}^2$. Show that the surface may be parametrised by setting

$$x = \frac{1+w^2}{2w}, \qquad y = \frac{1-w^2}{2iw} \text{ for } w \in \mathbf{C}, w \neq 0.$$

Find a homeomorphism from X to the cylinder

$$\{(x,y,z) : x^2 + y^2 = 1\}$$

in $\mathbf{R}^3$.

Exercise 92 Show that for every $\lambda \in \mathbf{R}$ the straight line with equations

$$x - z = \lambda(1 - y), \quad \lambda(x + z) = 1 + y$$

lies on the hyperboloid $x^2 + y^2 = 1 + z^2$. Find another family of lines on this hyperboloid and show that lines of the same family do not intersect, but that each line of the first family meets each line of the second.

Exercise 93 Let $V \cong \mathbf{R}^4$ be the space of polynomials $p(x) = x^2 + ax + b$ where $a, b \in \mathbf{C}$. Let $X \subset V$ be the set of polynomials with coincident roots.
(a) Find a chart $\mathbf{r} : \mathbf{R}^2 \to X$ for X.
(b) Let $\alpha \in \mathbf{C}$. Show that the tangent plane to X at $(x - \alpha)^2$ is precisely those polynomials $p \in V$ satisfying $p(\alpha) = 0$.

Exercise 94 Let $\{(x,y,0) : x, y \in \mathbf{R}\}$ be identified with $\mathbf{C}$ by setting $x + iy = (x,y,0)$. Let $N = (0,0,1)$ and let $\pi_N : S^2 - N \to \mathbf{C}$ denote stereographic projection from the north pole N. Let $P : \mathbf{C} \to \mathbf{C}$ be a polynomial with complex coefficients. Show that the map

$$\pi_N^{-1} \circ P \circ \pi_N : S^2 - N \to S^2 - N$$

is smooth. If we extend $\pi_N^{-1} \circ P \circ \pi_N$ to all of S^2 by setting,

$$\pi_N^{-1} \circ P \circ \pi_N(N) = N$$

is $\pi_N^{-1} \circ P \circ \pi_N$ smooth at N? How can we make sense of this question when N is not in the domain of π_N?

Exercise 95 When $\mathbf{C} \cup N$ is identified with S^2 as above show that the map $z \to 1/z$ is a diffeomorphism of S^2. Hence show that the Möbius transformation

$$z \mapsto \frac{az+b}{cz+d}, \qquad ad \neq bc,$$

is a diffeomorphism of S^2.

Exercise 96 (*) Let $f : \mathbf{R}^3 \to \mathbf{R}$ be a smooth function. We say that $p \in \mathbf{R}^3$ is a *critical point* of f if

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0$$

at p and $f(p)$ is said to be a *critical value* of f. Find the critical points and values of the following functions.

(a) $f(x,y,z) = x^2 + y^2 + z^2$,
(b) $f(x,y,z) = x^2 + y^2 - \cosh^2 z$,
(c) $f(x,y,z) = (x^2 + y^2)/(1 + z^2)$.

Given that if $f : \mathbf{R}^3 \to \mathbf{R}$ is a smooth function and $a \in \mathbf{R}$ is not a critical value of f then $f^{-1}(a)$ is a smooth surface in $\mathbf{R}^3$, (Do Carmo p. 59). Deduce that the sets

$$\{(x,y,z) \in \mathbf{R}^3 : x^2 + y^2 + z^2 = 1\},$$
$$\{(x,y,z) \in \mathbf{R}^3 : x^2 + y^2 = \cosh^2 z\},$$
$$\{(x,y,z) \in \mathbf{R}^3 : x^2 + y^2 - z^2 = 1\},$$

are all smooth surfaces.

The First Fundamental Form

Exercise 97 Find the area of the torus in $\mathbf{R}^3$ given by

$$\mathbf{r}(u,v) = ((a + b\cos v)\cos u, (a + b\cos v)\sin u, b\sin v)$$

for $u, v \in (0, 2\pi)$ and $a > b > 0$.

Exercise 98 Let S be a surface of revolution and C a generating curve. Let s denote arc-length on C and let $\rho(s)$ denote the distance from the axis of rotation to the point of C with parameter s. Show that the area of S equals

$$2\pi \int_0^l \rho(s)\,\mathrm{d}s$$

where l is the length of C.

Exercise 99 Show that a smooth surface of revolution may be parametrised so that

$$E = E(v), \qquad F = 0, \qquad G = 1.$$

Exercise 100 Let $\mathbf{r} : (u,v) \mapsto \mathbf{r}(u,v)$ be a smooth chart. Show that the solutions to the differential equation

$$A\dot{u}^2 + 2B\dot{u}\dot{v} + C\dot{v}^2 = 0,$$

all meet orthogonally if and only if

$$EC - 2FB + GA = 0.$$

Exercise 101 Let $\gamma : [a,b] \to \mathbf{R}^3$ be a curve parametrised by arc-length. For each $u \in [a,b]$ let Π_u denote the plane through $\gamma(u)$ normal to the curve, and let S denote the surface swept out, as u varies, by the circle in Π_u of centre $\gamma(u)$ and constant radius r.

(i) Explain why the surface S can be parametrised by

$$\mathbf{r}(u,v) = \gamma(u) + r(\mathbf{n}(u)\cos v + \mathbf{b}(u)\sin v),$$

where $\mathbf{n}$ and $\mathbf{b}$ are the unit normal and binormal vectors to γ.

(ii) Show that the unit normal vector to S at the point parametrised by (u, v) is

$$\mathbf{N}(u, v) = -(\mathbf{n}(u)\cos v + \mathbf{b}(u)\sin v).$$

(iii) Show that the first fundamental form of S is

$$((1 + \kappa r\cos v)^2 + r^2\tau^2)\mathrm{d}u^2 + 2r^2\tau\mathrm{d}u\mathrm{d}v + r^2\mathrm{d}v^2,$$

where κ and τ are the curvature and torsion of the original curve γ. Hence prove that the area of S is equal to $2\pi r$ times the length of γ.

Exercise 102 Two curves on the same smooth parametrised surface are given parametrically by $t \mapsto (u(t), v(t))$ and $t \mapsto (\tilde{u}(t), \tilde{v}(t))$. Suppose that the curves intersect at t=0. (i.e. $u(0) = \tilde{u}(0)$ and $v(0) = \tilde{v}(0)$.) Prove that the angle of intersection θ is given by

$$\cos\theta = \frac{E\dot{u}\dot{\tilde{u}} + F(\dot{u}\dot{\tilde{v}} + \dot{\tilde{u}}\dot{v}) + G\dot{v}\dot{\tilde{v}}}{\sqrt{(E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2)}\sqrt{(E\dot{\tilde{u}}^2 + 2F\dot{\tilde{u}}\dot{\tilde{v}} + G\dot{\tilde{v}}^2)}}$$

Deduce that a chart is conformal if and only if the first fundamental form satisfies $E = G$ and $F = 0$ everywhere.

Exercise 103 A diffeomorphism between surfaces X and Y is said to be *conformal* if the angle between any two intersecting curves on X equals the angle between their images on $Y$and is said to be *area-preserving* if each subset of X is mapped to a subset of Y of equal area. Show that a diffeomorphism is an isometry if and only if it is area-preserving and conformal.

Exercise 104 Show that the normal maps from both the catenoid

$$x^2 + y^2 = \cosh^2 z$$

and the helicoid

$$x\sin z = y\cos z$$

to the sphere are conformal.

Exercise 105 Let $G : \mathbf{R}^3 \to \mathbf{R}^3$ be a map such that

$$|G(p) - G(q)| = |p - q|$$

for all $p, q \in \mathbf{R}^3$. G is said to be a *distance preserving map.* Prove that there is a linear orthogonal map $F : \mathbf{R}^3 \to \mathbf{R}^3$ and $p_0 \in \mathbf{R}^3$ such that

$$G(p) = F(p) + p_0$$

for all $p \in \mathbf{R}^3$.

Let S be a surface in $\mathbf{R}^3$ such that $G(S) \subseteq S$. Show that the restriction of G to S is an isometry. Find an example of a surface S and an isometry $H : S \to S$ which is not the restriction to S of a distance preserving map $G : \mathbf{R}^3 \to \mathbf{R}^3$.

Curvature and the Weingarten Map

Exercise 106 Find the lines of curvature and the principal curvatures on a surface of revolution in terms of the distance ρ of the generating curve from the axis. Show that the Gaussian curvature K equals $\kappa\cos\phi/\rho$ where κ is the curvature of the generating curve and ϕ is the angle between the axis and the tangent line to the curve.

Exercise 107 Find the second fundamental form, the principal curvatures and the Gaussian curvature of the catenoid and helicoid (see Example 47). Show that the isometry between the catenoid and the helicoid in Example 47 leaves the value of the Gaussian curvature unchanged.

Exercise 108 A smooth curve C in the xz-plane is parametrised by arc-length by

$$(f(s), g(s)), \quad -\pi/2 \leq s \leq \pi/2,$$

and C meets the z-axis only orthogonally at $z = \pm 1$, when $s = \pm\pi/2$. Let S be the surface of revolution parametrised by

$$(f(s)\cos t, f(s)\sin t, g(s)).$$

Show that at any point of S with $s \neq \pm\pi/2$ the principal directions are tangent to the co-ordinate curves and the Gaussian curvature is

$$(\dot{f}\ddot{g} - \dot{g}\ddot{f})\dot{g}/f.$$

Show that for the unit sphere in $\mathbf{R}^3$ the Gaussian curvature is constant. Prove that if the Gaussian curvature of the above surface S is constant then S is the unit sphere.

Exercise 109 Suppose that a smooth surface Σ contains the origin and lies entirely in the half-space $\{(x, y, z) \in \mathbf{R}^3 : z \geq 0\}$. Show that, if the principal curvatures of Σ at the origin are not zero, then they have the same sign.

Now suppose that, for some positive R, the surface Σ is contained in the ball

$$\{(x, y, z) : x^2 + y^2 + (z - R)^2 = R^2\}.$$

Show that the absolute value of each principal curvature of Σ at the origin is not less than R^{-1}. Hence deduce that the Gaussian curvature of a compact smooth surface in $\mathbf{R}^3$ is positive at some point.

Exercise 110 Let X and Y be the surfaces in $\mathbf{R}^3$ given by

$$\mathbf{x}(u, v) = (u\cos v, u\sin v, v) \qquad u, v \in (0, 2\pi)$$
$$\mathbf{y}(u, v) = (v\sin u, v\cos u, \log v) \qquad u, v \in (0, 2\pi)$$

respectively. Find the principal curvatures for both surfaces. Show that there is a bijection between X and Y which preserves the Gaussian curvature K but which is not an isometry. (Gauss' *Theorema Egregium* states that Gaussian curvature is invariant under isometries. We can see from the above example that the converse of the Theorema Egregium does not hold.)

Exercise 111 A *ruled surface* X is a surface of the form $\mathbf{r}(u, v) = \boldsymbol{\alpha}(u) + v\mathbf{w}(u)$ where $\boldsymbol{\alpha}$ is a smooth curve in $\mathbf{R}^3$, known as the *directrix*, and $|\mathbf{w}(u)| = 1$ for all u. X is said to be *non-cylindrical* if $\mathbf{w}'(u) \neq \mathbf{0}$ for all u.
a) Let X be a non-cylindrical ruled surface as above. Show that there is a curve of the form $\boldsymbol{\beta}(u) = \boldsymbol{\alpha}(u) + v(u)\mathbf{w}(u)$ such that $\boldsymbol{\beta}' \cdot \mathbf{w}' = 0$. Show that $\boldsymbol{\beta}$ is independent of the choice of directrix. $\boldsymbol{\beta}$ is known as the *line of striction.*
b)Note that $\boldsymbol{\beta}'(u) \wedge \mathbf{w}(u) = \lambda(u)\mathbf{w}'(u)$ for some $\lambda(u)$. By taking the line of striction as directrix show that any singular points of the ruled surface (i.e. where $\mathbf{r}_u \wedge \mathbf{r}_v = \mathbf{0}$) lie on the line of striction.
c) Show that the Gaussian curvature is given by $K = -\lambda^2(\lambda^2 + v^2)^{-2}$.

Exercise 112 (*) Let X be a smooth compact surface in $\mathbf{R}^3$, let S^2 denote the sphere of radius one with centre the origin in $\mathbf{R}^3$, and let $\mathbf{N} : X \to S^2$ denote the mapping defined by assigning to each point of X the outward-pointing unit normal vector to X. Suppose that the restriction of $\mathbf{N}$ to a connected open subset U of X is a bijection onto $\mathbf{N}(U)$ and that the Gaussian curvature K is nowhere zero on U.

(a) Show that the area of $\mathbf{N}(U)$ equals the absolute value of $\int_U K \, \mathrm{d}A$.

(b) Deduce that if X is a torus of revolution then $\int_X K \, \mathrm{d}A = 0$.

Exercise 113 (*) Determine explicitly the smooth functions ρ for which the surface of revolution

$$\mathbf{r}(u, v) = (\rho(u)\cos v, \rho(u)\sin v, u)$$

has mean curvature identically zero.

Exercise 114 Let X be a smooth parametrised surface, γ a smooth curve in X and $\mathbf{t}_p$ the unit tangent to γ at $p \in \gamma$. Then γ is said to be *asymptotic* if $W_p \mathbf{t}_p \cdot \mathbf{t}_p = 0$ for each $p \in \gamma$. Show that the normal $\mathbf{n}_p$ of an asymptotic curve (where defined) is tangent to X. By expressing W_p as a matrix relative to the basis $\{\mathbf{t}_p, \mathbf{n}_p\}$ for T_pX, or otherwise, deduce that the torsion τ of an asymptotic curve is given by $\tau^2 = -K$.

Exercise 115 Show that on the helicoid S parametrised by

$$(s\cos u, s\sin u, u) \qquad s, u \in \mathbf{R},$$

each curve C_k given by $s = k$ (where k is constant) is asymptotic.

Now suppose that a smooth curve C in the xy-plane is parametrised by arc-length by

$$(f(s), g(s)) \quad 0 \leq s \leq 1,$$

and that C meets the y-axis only when $s = 0$. Consider the surface S' parametrised by

$$(f(s)\cos u, f(s)\sin u, g(s) + u), \quad 0 < s < 1, \quad u \in \mathbf{R}.$$

Prove that each curve C'_k on S' given by $s = k$ (where k is constant, $0 < k < 1$) is asymptotic then S' is isometric to part of S.

Exercise 116 (*) Let Σ be a surface in $\mathbf{R}^3$ whose Gaussian curvature is zero everywhere, but whose mean curvature is nowhere zero. You may assume that in the neighbourhood of any point of Σ there is a parametrisation $\mathbf{r}(u, v)$ for Σ such that the first and second fundamental forms can be expressed as

$$Edu^2 + Gdv^2 \text{ and } Ldu^2,$$

respectively. Show that the normal $\mathbf{n}$ and its derivative $\dfrac{\partial \mathbf{n}}{\partial u}$ are independent of the co-ordinate v. By considering the vector product $\mathbf{n} \times \dfrac{\partial \mathbf{n}}{\partial u}$, or otherwise, show that

$$\frac{\partial}{\partial v}(\psi \frac{\partial \mathbf{r}}{\partial v}) = 0,$$

for some non-vanishing function ψ, and deduce that Σ is ruled (that is, each point of Σ is contained in a straight line segment which lies in Σ). Let $\gamma : \mathbf{R} \to \mathbf{R}^3$ be a smooth curve with $|\gamma'(t)| = 1$, and such that γ'' is nowhere zero. Let Σ be the surface with parametrisation

$$\mathbf{r}(t, v) = \gamma(t) + v\gamma'(t).$$

Show that the first fundamental form of Σ is

$$ds^2 = dv^2 + 2dvdt + (1 + v^2|\gamma''(t)|^2)dt^2.$$

By considering a suitable plane curve, show that there is an open set in Σ which is isometric to an open set in $\mathbf{R}^2$.

Exercise 117 (*) Let U be the rectangle

$$\{(u, v) : -\pi \leq u \leq \pi, 0 < v < 2\}$$

endowed with the first fundamental form

$$(1 + v^2)\mathrm{d}u^2 + \mathrm{d}v^2.$$

The surface X is then formed by identifying the point $(-\pi, v)$ with (π, v) for each $0 < v < 2$.

Decide which of the following statements about X are true, which false. Justify your answers.
(a) The map $\mathbf{r} : X \to \mathbf{R}^3$ given by

$$\mathbf{r}(u, v) = (\sqrt{1 + v^2}\cos u,\ \sqrt{1 + v^2}\sin u,\ \sinh^{-1} v)$$

for $-\pi \leq u \leq \pi$, $0 < v < 2$ is an isometry from X onto $\mathbf{r}(X)$.
(b) The curves given by $v = 1 + u$ and $v = 1 - u$ meet at right angles in X.
(c) Let $0 \leq k < 2\pi$. The maps ϕ_k and ψ, defined by:

$$\phi_k(u,v) = \begin{cases} (u+k, v) & -\pi \leq u \leq \pi - k, \\ (u+k-2\pi, v) & \pi - k \leq u \leq \pi, \end{cases}$$

and

$$\psi(u,v) = (-u, v),$$

are isometries of X.

Geodesics

Exercise 118 Show that a curve of constant geodesic curvature c on the unit sphere is a planar circle of length $2\pi/\sqrt{1+c^2}$.

Exercise 119 Let $\mathbf{r} : (u,v) \mapsto \mathbf{r}(u,v)$ be a smooth chart. Write each of the following in terms of the coefficients of the first fundamental form E, F, G and their derivatives:

$$\mathbf{r}_{uu} \cdot \mathbf{r}_u, \quad \mathbf{r}_{uv} \cdot \mathbf{r}_u, \quad \mathbf{r}_{vv} \cdot \mathbf{r}_u, \quad \mathbf{r}_{uu} \cdot \mathbf{r}_v, \quad \mathbf{r}_{uv} \cdot \mathbf{r}_v, \quad \mathbf{r}_{vv} \cdot \mathbf{r}_v.$$

Exercise 120 Let U be an open subset of $\mathbf{R}^2$ and let $\mathbf{r} : U \to \mathbf{R}^3$ be a parametrisation of $\mathbf{r}(U)$, a smooth patch of surface in $\mathbf{R}^3$. Let γ be a smooth curve in $\mathbf{r}(U)$ parametrised by arc-length and let

$$\mathbf{n} = \frac{\mathbf{r}_u \wedge \mathbf{r}_v}{|\mathbf{r}_u \wedge \mathbf{r}_v|}$$

be a choice of unit normal. The *geodesic curvature* of γ is defined to be

$$k_g = \ddot{\gamma} \cdot (\mathbf{n} \wedge \dot{\gamma}).$$

Suppose now that the first fundamental form of $\mathbf{r}(U)$ is $E\mathrm{d}u^2 + G\mathrm{d}v^2$. Prove that

$$\mathbf{n} \wedge \dot{\gamma} = \dot{u}\sqrt{\frac{E}{G}}\mathbf{r}_v - \dot{v}\sqrt{\frac{G}{E}}\mathbf{r}_u.$$

Show further that the geodesic curvature of γ equals

$$\sqrt{EG}(\dot{u}\ddot{v} - \dot{v}\ddot{u}) + P\dot{u}^3 + Q\dot{u}^2\dot{v} + R\dot{u}\dot{v}^2 + S\dot{v}^3$$

where P, Q, R, S may be written in terms of E, G and their derivatives and write P explicitly in this way. (You may assume, without proof, the result of the previous exercise.)

Exercise 121 Write down the differential equations determining the geodesics on the graph $z = f(x,y)$.

Exercise 122 Define what is meant by a *geodesic* on a surface $X \subset \mathbf{R}^3$. The curve $\mathbf{r}(t) = (x(t), y(t), z(t))$ lies on the surface X given by the equation $f(x,y,z) = 0$ where $\partial f/\partial x, \partial f/\partial y$, and $\partial f/\partial z$ vanish nowhere on X. Show that the curve is a geodesic if and only if

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} \Big/ \frac{\partial f}{\partial x} = \frac{\mathrm{d}^2y}{\mathrm{d}t^2} \Big/ \frac{\partial f}{\partial y} = \frac{\mathrm{d}^2z}{\mathrm{d}t^2} \Big/ \frac{\partial f}{\partial z}.$$

Show that the surface X is a surface of revolution if and only if there is a line L such that all planes through L meet X in geodesics.

Exercise 123 Let S denote the surface $x^2 + y^2 - z^2 = 1$ and let γ be a geodesic in S. Then $h = \rho \sin \psi$ is constant by Clairaut's relation where $\rho(t)$ is the distance from $\gamma(t)$ to the axis of rotation and $\psi(t)$ is the angle between $\gamma'(t)$ and the meridian through $\gamma(t)$.

Show that if $|h| > 1$ then γ remains in one of the half-spaces $z > 0$ or $z < 0$, while if $|h| < 1$ γ passes through the $z = 0$ plane. What happens if $h = \pm 1$?

Exercise 124 A surface S has first fundamental form

$$E(u,v) = G(u,v) = U(u) + V(v), \qquad F = 0.$$

Show that on a geodesic in S

$$U \sin^2\theta - V\cos^2\theta$$

is constant, where θ is the angle the geodesic makes with the curve $v =$ constant. Explain how this is a generalisation of Clairaut's relation.

Exercise 125 Starting from any property which characterises a geodesic, obtain the differential equations satisfied by the geodesics on a surface with the metric

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2.$$

Show that the curves $u =$ constant are all geodesics if and only if

$$GG_u + FG_v = 2GF_v$$

for all u, v.

Hence or otherwise prove that on the surface of revolution

$$x = f(u)\cos v, \qquad y = f(u)\cos v, \qquad z = u,$$

the curves $v =$ constant are always geodesics but the circles $u =$ constant are all geodesics if and only if the surface is a cylinder.

Exercise 126 (*) Show that the geodesic equations of the plane when endowed with the first fundamental form $du^2 + e^{2u}dv^2$ may be written as

$$\ddot{u} = c\dot{v}, \qquad \dot{v}^2 = ce^{-2u},$$

for some constant $c \geq 0$, and hence find the geodesics of this first fundamental form.

Show that the map from $\mathbf{R}^2$ to the upper half plane given by $\phi : (u,v) \mapsto (v, e^{-u})$ is a diffeomorphism. For what first fundamental form on the upper half plane is ϕ an isometry?

Further Topics

Exercise 127 (*) Let Σ denote the tractoid (see Example 44) with one meridian removed. Using the fact that Σ is isometric to part of the hyperbolic plane show that for each point P of Σ there is a geodesic segment which begins at P and returns to P after passing once around Σ. Show that if $P = (x(t), y(t), 0)$ the angle between the initial and final tangent vectors of the geodesic segment is $2\tan^{-1}(\pi \sin t)$.

Exercise 128 (*) The hyperbolic plane can be defined as the unit disc $D = \{z \in \mathbf{C} : |z| < 1\}$ endowed with the first fundamental form

$$\frac{4|dz|^2}{(1-|z|^2)^2}$$

where $z = x + iy$ and $|dz|^2 = dx^2 + dy^2$. Write down the equations which must be satisfied in order for a curve parametrised by arc-length to be a geodesic in D, and verify that the real axis with the parametrisation $x = \tanh(s/2) \in (-1,1), y = 0$, is one.

For each $a \in D$ and $\theta \in [0, 2\pi)$, define $f_{a,\theta}$ on D by

$$f_{a,\theta}(z) = e^{i\theta}\frac{z-a}{1-\bar{a}z}$$

and show that $f_{a,\theta}$ is an isometry of D onto itself. Deduce, or prove otherwise, that any two points $a, b \in D$ lie on a geodesic, along which they are separated by arc-length

$$2\tanh^{-1}\left|\frac{b-a}{1-\bar{a}b}\right|.$$

Exercise 129 (*) Show that the function

$$d(a,b) = 2\tanh^{-1}\left|\frac{b-a}{1-\overline{a}b}\right|$$

for $a, b \in D$ defines a metric on D.

Exercise 130 (*) Prove that a hyperbolic circle is simply an ordinary circle in D, but that its hyperbolic centre is usually not its ordinary centre. Prove that a hyperbolic circle of radius a has circumference $2\pi\sinh a$ and area $4\pi\sinh^2(a/2)$.

Exercise 131 (*) The real projective plane $\mathbf{P}$ is the quotient space

$$\mathbf{P} = \frac{\mathbf{R}^3 - \{\mathbf{0}\}}{\mathbf{R} - \{0\}}.$$

Let $[x : y : z]$ denote the homogeneous co-ordinate of the line through $\mathbf{0}$ and (x, y, z); i.e. the ratio $x : y : z$ common to all points on that line. Show that the charts $\phi_i : \mathbf{R}^2 \to \mathbf{P}$ defined by

$$\phi_1(x,y) = [1 : x : y], \quad \phi_2(x,y) = [x : 1 : y], \quad \phi_3(x,y) = [x : y : 1],$$

form a smooth atlas for $\mathbf{P}$. Are the transition maps of this atlas orientation-preserving?

The map $F : \mathbf{R}^3 - \{\mathbf{0}\} \to \mathbf{R}^4$ is defined by

$$F(x,y,z) = \frac{1}{(x^2+y^2+z^2)}(\frac{1}{2}(x^2-y^2), xy, xz, yz).$$

Show that F induces a map $\tilde{F}$ on $\mathbf{P}$. This map is an embedding of $\mathbf{P}$ in $\mathbf{R}^4$. [You are not asked to prove this.] Show that the area of $\tilde{F}(\mathbf{P})$ is less than 2π. Let π denote the natural map from the unit sphere S^2 to $\mathbf{P}$. Is $\tilde{F} \circ \pi$ a local isometry?

Exercise 132 (*) Let S^2 denote the unit sphere. The map $f : \mathbf{R}^3 \to \mathbf{R}^5$ is given by

$$f(x,y,z) = (yz, zx, xy, \frac{1}{2}(x^2-y^2), \frac{1}{2\sqrt{3}}(1-3z^2)).$$

(i) Show that $f(\mathbf{p}) = f(\mathbf{q})$ for $\mathbf{p}, \mathbf{q} \in \mathbf{R}^3$ if and only if $\mathbf{p} = \pm\mathbf{q}$.
(ii) Show that the restriction of f to S^2 is a local isometry. So f induces an isometric embedding of the elliptic plane in $\mathbf{R}^5$.

Exercise 133 (*) Let $f : \mathbf{R}^m \to \mathbf{R}^n$ be a function and $x \in \mathbf{R}^m$. Then f is said to be differentiable at x if there is a linear map $Df(x) : \mathbf{R}^m \to \mathbf{R}^n$ such that

$$f(x+h) = f(x) + Df(x)(h) + o(h)$$

for all vectors $h \in \mathbf{R}^m$. ($o(h)$ represents any function $r : \mathbf{R}^m \to \mathbf{R}^n$ such that $|r(h)|/|h| \to 0$ as $h \to 0$.)
(i) Show that $Df(x)$ may be represented by a matrix whose entries are the partial derivatives of the components of the map f.
(ii) Prove the *chain rule*. That is, show for differentiable maps $f : \mathbf{R}^m \to \mathbf{R}^n$ and $g : \mathbf{R}^n \to \mathbf{R}^p$, that

$$D(g \circ f)(x) = Dg(f(x)) \circ Df(x).$$