NTNU
Norwegian University of
Science and Technology

Geodesics on Surfaces

Eirik Ingebrigtsen

Master of Science in Physics and Mathematics
Submission date: June 2016
Supervisor: Lars Peter Lindqvist, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Abstract

We study geodesics on surfaces in the setting of classical differential geometry. We define the curvature of curves and surfaces in three-space and use the fundamental forms of a surface to measure lengths, angles, and areas. We follow Riemann and adopt a more abstract approach, and use tensor notation to discuss Gaussian curvature, Gauss's Theorema Egregium, geodesic curves, and the Gauss-Bonnet theorem. Properties of geodesics are proven by variational methods, showing the connection between straightest and shortest for curves on surfaces. The notion of intrinsic and extrinsic properties is highlighted throughout.

Sammendrag

Vi studerer geodetiske kurver på flater i klassisk differensialgeometri. Vi definerer krumning til kurver og flater i rommet, og bruker en flates fundamentalformer til å måle lenger, vinkler og areal. Vi følger Riemann i hans mer abstrakte tilnærming og bruker tensornotasjon i vår diskusjon av Gauss-krumningen, Gauss' Theorema Egregium, geodetiske kurver og Gauss-Bonnet-teoremet. Sammenhengen mellom korteste vei på flater og de geodetiske kurvene demonstreres ved variasjonsregning. Forskjellen mellom intrinsiske og ekstrinsiske egenskaper står sentralt.

Preface

This master thesis was written during my final five months at the Norwegian University of Science and Technology in Trondheim. Its submission marks the completion of my five-year long integrated M.Sc. in Applied Physics and Mathematics, with specialisation Industrial Mathematics.

Before working on differential geometry, as is the underlying topic of this thesis, I had followed a variety of courses spanning from introductory physics, via management and programming to engineering mathematics. Amongst the mathematical topics I had studied, statistics and numerical analysis had been the cornerstones. As the applied mathematics concepts got more and more involved, I felt the need to reinforce my foundations by studying pure mathematics.

It was at this point I ended up focusing on the historically important calculus of variation. This concept had been introduced in a course on optimisation and in connection with the finite element method, but without much rigour. Professor Peter Lindqvist helped me get started working on this theory. In the autumn, I focused on variational methods for PDEs, in particular, the Dirichlet eigenvalues, while attending courses in optimal control theory and convex analysis.

Choosing a final thesis topic was not easy. Part of me wanted to dig deeper in the theory of PDEs, while another was curious as to what other beautiful mathematical concepts I had missed. The latter side of me won when Professor Peter Lindqvist pitched the idea of geodesics and the geometry of surfaces. I embraced my 'Jack of all trades, master of none'-side and started from scratch, refreshing results from multivariate calculus. I had not worked on differential geometry before, so this was my first introduction to the subject.

I would like to thank Professor Peter Lindqvist for allowing me to work on established theory, for proposing an interesting topic, and for guiding me safely the last year. I would also like to thank Lene and my family for their support.

Eirik Ingebrigtsen
Trondheim, June 2016

Contents

Chapter 1

Introduction

What can a mathematically inclined two-dimensional being living on a surface figure out about the geometry of its world? What are the intrinsic properties of a surface? These questions lay at the centre of a major breaking point in history, where mathematicians and physicists questioned the structure of space itself. We will in this thesis develop the theory of curves and surfaces in Euclidean space with the goal of finding the tools to answer such questions. The key one will be the ‘straight’ lines of curved space: the geodesics.

Central to the discussion is the distinction between intrinsic and extrinsic properties of geometric objects. An intrinsic property is something that does not depend on how, and in which context, we represent the object, while an extrinsic property does. Imagine that you are holding an inelastic rope. No matter how you hold the rope in space, the length stays the same. Thus, the length is an intrinsic property. The curvature, however, will depend on how we hold it. Keeping it straight gives the rope zero curvature, while coiling it up gives it a lot, showing that this indeed is an extrinsic property.

We now outline the history of geometry up to the time of Bernhard Riemann. This treatment is in no way complete and any book on the history of mathematics will do a better job, e.g. the in-depth treatment, on which this introduction is based, found in the recent book *5000 Years of Geometry* [1].

Geometry has always fascinated the mathematician. Even the earliest mathematical texts focus on geometric problems. The mathematicians and philosophers of ancient Greece made substantial contributions to science, in particular with the formal concept of ‘proof’. With Euclid’s *Elements* came the most influential work on geometry of all time. These books became the basis of mathematical education all the way to the mid-20th century [2]. The Euclidean geometry[1] is based on an axiomatic system, a set of true statements about geometry. These appealing postulates allowed mathematicians to prove a wealth of propositions concerning geometric objects.

A first leap forward came when René Descartes and Pierre de Fermat started using algebraic techniques to study geometric problems in the early 1600s. This new analytic geometry, where reference to a coordinate system is central, allowed for new reasoning with the axioms of algebra, not those of geometry. In this paradigm, the relationship between algebraic equations

[1]The term Euclidean is of course added centuries later, as it at the time was the only form of geometry.

and curves and surfaces was extensively studied. After the introduction of analytic geometry and Cartesian coordinates, the approach to geometry without any coordinates was named synthetic geometry. Both methods are in use today.

The first systematic treatment of the curvature of surfaces came with Euler in 1767 [3]. At the time, they knew how to describe the curvature of a curve. They used the osculating circle, which, informally, is the circle that best matches the curve at a point (see section 2.1). Euler points out that, although the curvature of a sphere should be that of a great circle, using the sphere that locally best 'matched' a surface is inadequate, e.g. for a cylinder or a saddle. Using cross sections, he arrives at what today is called the principal curvatures (see section 3.3).

The first notion of intrinsic properties came with Gauss [4], who also unified the theory of surfaces. He studied what it meant for a curve on a surface to be 'straight', and found this to be a helpful tool. We will, for the most part, follow in his footsteps, and in particular prove his remarkable theorem showing that the curvature of a surface is intrinsic. This has a direct application to cartography, as it shows why it is impossible to make a flat map of Earth that does not distort at least some distances. Cartography and geodesy were up to modern time the main forces behind the development of geometry.

> 'That, if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles [5].'

This is a translation of Euclid's fifth postulate, known as the parallel postulate. The parallel postulate has intrigued many mathematicians over the years. It was not as self-evident as the rest of Euclid's postulates, and many tried to prove it from the other four. A geometry where the fifth postulate does not hold is today known as a non-Euclidean geometry. The geometries of the general surfaces of Gauss and his student Riemann are such examples. As are the geometries of János Bolyai[2] and Lobachevsky, who, at roughly the same time, also developed new geometries. More details on the geometry of Riemann will follow in chapter 6. The development of geometry has since then exploded, with the most popular application being the sub-Riemmanian geometry of Einstein's general relativity.

The treatment presented here is close to the classical treatment of differential geometry, so for a reader in the know; we do not talk about the shape operator, manifolds nor connections. We will, however, follow the works of Gauss, and try to answer questions such as: How can we describe curves on a surface? What are the straight lines of curved space? What can be found

[2] The history of father and son Bolyai's dispute with Gauss is worth seeking out.

out by intrinsic reasoning only? Can a two-dimensional being living on a surface figure out the geometry of its world?

Chapter 2 lay out the elementary theory of curves and surfaces in three-space. We define what we mean by a curve, and a curve's arc length, speed, curvature and torsion, and prove the fundamental theorem of space curves. Then, we take a modern definition of a smooth surface by surface patches, homeomorphisms, and transition maps. Describing curves on surfaces then helps us define the intuitive version of a tangent space and normals to a surface. The section ends with the first fundamental form, a tool that allows us to measure lengths, angles and areas.

Chapter 3 is dedicated to the description of the curvature of surfaces. After introducing tensor-notation, we connect the geometric structure of a surface to the curves on the surface. We define the second fundamental form and the different curvatures of a surface. Finally, the historically significant Theorema Egregium is presented.

Chapter 4 concerns the geodesics on a surface, i.e. the curves that do not bend relative to the surface. We show that the Christoffel symbols used in the proof of Theorema Egregium depend only on the first fundamental form, completing the proof of the Theorema Egregium. Then we describe the geodesic curvature and show the geodesic equations, and connect the geodesic equations to the length-minimising curves on a surface. The section ends by defining geodesic coordinate systems and proving the Gauss-Bonnet theorem.

Chapter 5 is a worked example trying to show that, even for an elementary surface, finding all geodesics is typically hard. The surface chosen is the catenoid, a surface of revolution that also serves as the prototype of a minimal surface.

Chapter 6 is devoted to bringing the main results of section 3 and 4 into the setting of Riemannian geometry. Here, we discuss the impact of Riemann's work and explain how his abstractions sparked the modern era of geometry.

All vector graphics are made with the free editor Inkscape [6], while the surface plots are generated in Matlab [7]. No code is included.

Chapter 2

Curves and Surfaces

This introductory chapter will introduce and investigate real curves and smooth surfaces in three-dimensional Euclidean space. The concepts presented here are elementary, but integral to formulating and understanding the straight lines of curved space: the geodesics. It is an attempt to keep the discussion short, while still being self-contained. For a full treatment any introductory book on differential geometry will do. For a first reading I would recommend the book *Elementary Differential Geometry* by A. N. Pressley [8]. The notation used is an adoption of the notation found in that book, in combination with that found in *Differential Geometry and Relativity Theory* by R. L. Faber [9] and *Differential Geometry* by E. Kreyszig [10].

2.1 Curves in $\mathbb{R}^3$

In this section we introduce the terminology of space curves, derive the Serret-Frenet formulas and prove the fundamental theorem of space curves. The basic idea of a curve is shown in Figure 2.1.

Curves in $\mathbb{R}^3$ may be seen as paths traced out by a moving point. If $\boldsymbol{\alpha}(t)$ is the position vector at time t, the curve is described by the vector-valued function $\boldsymbol{\alpha}$ of the scalar parameter t.

Definition 1 (Curve). A smooth parameterised curve in $\mathbb{R}^3$ is a map $\boldsymbol{\alpha} : (a, b) \to \mathbb{R}^3$

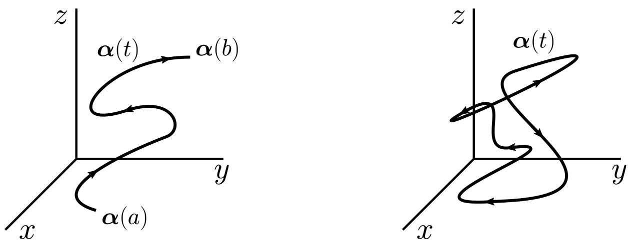
$$\boldsymbol{\alpha}(t) = (x(t), y(t), z(t)) \quad t \in (a, b) \tag{2.1.1}$$

where $-\infty \leq a \leq b \leq \infty$ and each component is a smooth function of t.

In this text, all curves are assumed smooth in the sense that the coordinate functions are sufficiently differentiable on the whole domain. The reason for this is to avoid having details of differentiability obstruct the main discussion. Further, all curves are assumed regular, i.e. its derivative never vanishes. Or more intuitively, the moving point never come to a complete stop or backtracks.

Definition 2 (Velocity vector). The velocity vector $d\boldsymbol{\alpha}/dt = \boldsymbol{\alpha}'$ is tangent to the curve, points in the direction of increasing t, and is given by

$$\boldsymbol{\alpha}'(t) = (x'(t), y'(t), z'(t)).$$



(a) (b)

Figure 2.1: The figure depicts two curves in $\mathbb{R}^3$. The small arrows indicate the direction of increasing t. (b) is an example of a closed curve, i.e. a curve for which there is a constant $c > 0$ such that $\boldsymbol{\alpha}(t + c) = \boldsymbol{\alpha}(t)$ for all t.

Definition 3 (Arc length). The arc length $s = s(t)$ of a curve $\boldsymbol{\alpha}$ measures the length of the curve from initial point $\boldsymbol{\alpha}(a)$ to the point $\boldsymbol{\alpha}(t)$ and is given by

$$s(t) = \int_a^t \|\boldsymbol{\alpha}'(u)\| du, \tag{2.1.2}$$

where $\| \cdot \|$ denotes the standard Euclidean norm.

With s being the arc length,

$$\frac{ds}{dt} = \frac{d}{dt} \int_a^t \|\boldsymbol{\alpha}'(u)\| du = \|\boldsymbol{\alpha}'(t)\|$$

by the fundamental theorem of calculus, and so ds/dt can be seen as the speed of the curve at $\boldsymbol{\alpha}(t)$.

Definition 4 (Unit-speed curve). If $\boldsymbol{\alpha}$ is a curve, its speed at $\boldsymbol{\alpha}(t)$ is $\|\boldsymbol{\alpha}'(t)\|$ and $\boldsymbol{\alpha}$ is a unit-speed curve if $\boldsymbol{\alpha}'(t)$ is a unit vector for all $t \in (a, b)$.

One can obviously have different parameterisations of the same curve. We therefore say that two smooth curves $\boldsymbol{\alpha}_1 : I_1 \to \mathbb{R}^3$ and $\boldsymbol{\alpha}_2 : I_2 \to \mathbb{R}^3$ are *equivalent* if there is a smooth bijective map $\varphi : I_2 \to I_1$ whose inverse $\varphi^{-1} : I_1 \to I_2$ is also smooth and $\boldsymbol{\alpha}_2(t) = \boldsymbol{\alpha}_1(\varphi(t))$ for all t. If such a map exist, $\boldsymbol{\alpha}_2$ is a reparameterisation of $\boldsymbol{\alpha}_1$. The most useful choice of parameter is $t = s$ where s is arc length. This gives $ds/dt = 1$ so that $\boldsymbol{\alpha}(s)$ is a unit-speed curve. We call $\boldsymbol{\alpha}'(s)$ the *unit tangent vector* and define

$$\mathbf{t}(s) = \boldsymbol{\alpha}'(s). \tag{2.1.3}$$

The rate of change of this vector of constant length is a measure of the curvature of $\boldsymbol{\alpha}$. The faster $\mathbf{t}$ turns, the larger the components of $\mathbf{t}'$ are, and thus we take the following definition.

Definition 5 (Curvature). The curvature of $\boldsymbol{\alpha}$ at $\boldsymbol{\alpha}(s)$ is the length of $\mathbf{t}'(s)$. We denote it by $k(s)$:

$$k(s) = \|\mathbf{t}'(s)\| = \|\boldsymbol{\alpha}''(s)\|.$$

Another way of thinking is that the curvature measures how much a curve differs from being a straight line in $\mathbb{R}^3$, since a curve being a straight line has vanishing curvature.

Lemma 1. *For a unit-speed curve $\boldsymbol{\alpha} : (a, b) \to \mathbb{R}^3$ the following three conditions are equivalent:*

(i) $k \equiv 0$,

(ii) $\boldsymbol{\alpha}'' \equiv 0$,

(iii) $\boldsymbol{\alpha}$ *is a straight line segment.*

Proof. By the definition of k we have that $(i) \Leftrightarrow (ii)$. Let us show that $(ii) \Leftrightarrow (iii)$. Assuming $\boldsymbol{\alpha}'' \equiv 0$, integration yields $\boldsymbol{\alpha}'(s) = \mathbf{u}$ where $\|\mathbf{u}\| = 1$ and then $\boldsymbol{\alpha}(s) = \mathbf{u}s + \mathbf{v}$ for the vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$. For $s \in (a, b)$ this is a unit-speed parameterisation of a straight line segment. Conversely, any unit-speed straight line can be given on the form $\boldsymbol{\alpha}(s) = \mathbf{u}s + \mathbf{v}$ for some $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ where $\|\mathbf{u}\| = 1$. Differentiation gives $\boldsymbol{\alpha}'' \equiv 0$. □

The unit vector in the direction of $\mathbf{t}'(s) = \boldsymbol{\alpha}''(s)$ is called the *principal normal vector* and is denoted by $\mathbf{n}(s)$. Hence we have

$$\mathbf{n}(s) = \frac{\mathbf{t}'(s)}{\|\mathbf{t}'(s)\|} = \frac{\mathbf{t}'(s)}{k(s)}$$

and

$$\mathbf{t}'(s) = k(s)\mathbf{n}(s).$$

Where $k(s) \neq 0$, the curve may, at the point $\boldsymbol{\alpha}(s)$, be approximated by the circle of radius $1/k(s)$ that is tangent to the curve and lies in the plane spanned by $\mathbf{t}(s)$ and $\mathbf{n}(s)$. This circle is called the *osculating circle*, and it shares curvature, tangent vector and principal normal vector with the curve at $\boldsymbol{\alpha}(s)$. Figure 2.2 shows the idea. This circle is said to lay in the *osculating plane*[1]. The unit normal to this plane is called the *binormal vector* and is given by the cross product

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

The rate of change in this vector-valued function, i.e. $\mathbf{b}'(s)$, gives information on how much the osculating plane turns and tilts. We find, omitting (s) for readability,

$$\mathbf{b}' = (\mathbf{t} \times \mathbf{n})' = \mathbf{t}' \times \mathbf{n} + \mathbf{t} \times \mathbf{n}' = \mathbf{t} \times \mathbf{n}',$$

[1] The term osculating plane was first used by Tinseau in 1780, according to E. Kreysizg on page 33 of his book *Differential Geometry* [10]

where the last equality is due to $\mathbf{t}' = k\mathbf{n}$ and $\mathbf{n} \times \mathbf{n} = 0$. This shows that $\mathbf{b}'$ is orthogonal to both $\mathbf{t}$ and $\mathbf{n}'$ and so it must be a multiple of $\mathbf{n}$.

Definition 6 (Torsion). We define the torsion of $\boldsymbol{\alpha}$ to be the function $\tau = \tau(s)$ defined by the equation

$$\mathbf{b}' = -\tau \mathbf{n}.$$

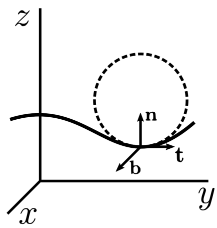


Figure 2.2: A curve living in the yz-plane, together with its osculating circle, unit tangent vector $\mathbf{t}$, principal normal vector $\mathbf{n}$, and binormal vector $\mathbf{b}$ at a point.

Example 1 (A circular helix). A circular helix can be described by the curve

$$\boldsymbol{\beta}(t) = (a\cos t, a\sin t, bt), \quad t \in \mathbb{R},$$

where $a > 0$ and b are constants. Let us compute a few of the quantities defined above. First let us find the speed $\|\boldsymbol{\beta}'(t)\|$, and to this end compute

$$\boldsymbol{\beta}'(t) = (-a\sin t, a\cos t, b)$$

and

$$\|\boldsymbol{\beta}'(t)\| = \left((-a\sin t)^2 + (a\cos t)^2 + b^2\right)^{1/2} = \sqrt{a^2 + b^2},$$

showing that the speed is constant. One turn of the helix, e.g. corresponding to t going from 0 to 2π, is

$$L = \int_0^{2\pi} \|\boldsymbol{\beta}'(t)\| dt = 2\pi\sqrt{a^2 + b^2}.$$

Let us reparameterise in terms of arc length. Take $s = t\sqrt{a^2 + b^2}$ so that $ds/dt = \|\boldsymbol{\beta}'(t)\|$. This gives the curve

$$\boldsymbol{\alpha}(s) = \left(a\cos\frac{s}{\sqrt{a^2 + b^2}}, a\sin\frac{s}{\sqrt{a^2 + b^2}}, \frac{bs}{\sqrt{a^2 + b^2}}\right)$$

and the unit tangent vector

$$\mathbf{t}(s) = \boldsymbol{\alpha}'(s) = \frac{1}{\sqrt{a^2 + b^2}}\left(-a\sin\frac{s}{\sqrt{a^2 + b^2}}, a\cos\frac{s}{\sqrt{a^2 + b^2}}, b\right).$$

Taking the derivative

$$\mathbf{t}'(s) = k(s)\mathbf{n}(s) = \frac{a}{a^2+b^2}\left(-\cos\frac{s}{\sqrt{a^2+b^2}}, -\sin\frac{s}{\sqrt{a^2+b^2}}, 0\right)$$

reveals the curvature

$$k(s) = \frac{a}{a^2+b^2}$$

and principle normal vector

$$\mathbf{n}(s) = \left(-\cos\frac{s}{\sqrt{a^2+b^2}}, -\sin\frac{s}{\sqrt{a^2+b^2}}, 0\right).$$

In terms of t this vector is $\mathbf{n}(t) = -(\cos t, \sin t, 0)$ and so the principle normal vector points inwards from the curve to the z-axis. To find the torsion, we first compute the binormal vector

$$\mathbf{b} = \mathbf{t} \times \mathbf{n} = \frac{1}{\sqrt{a^2+b^2}}\left(b\sin\frac{s}{\sqrt{a^2+b^2}}, -b\cos\frac{s}{\sqrt{a^2+b^2}}, a\right)$$

and its derivative

$$\mathbf{b}' = \frac{b}{a^2+b^2}\left(\cos\frac{s}{\sqrt{a^2+b^2}}, \sin\frac{s}{\sqrt{a^2+b^2}}, 0\right) = \tau\mathbf{n}.$$

This gives the torsion

$$\tau = \frac{b}{a^2+b^2}.$$

The constant torsion for the circular helix is due to the fact that, for a particle following the curve, the osculating plane turns at a constant rate.♣

The choice of parametric representation of a curve is of minor interest, but that does not indicate that the point set traced out is what defines the curve. Consider the two examples taken from Section 7 in [10] and shown in Figure 2.3.

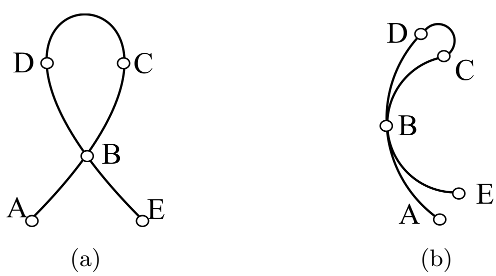


Figure 2.3: Two point sets whose geometric properties depend on the order of traversal.

In 2.3a we see that according to our previous definition, we would not get a smooth curve if we traverse the point set in the order $ABDCBE$. The reason for this being that at the point B the first derivative of the vector function does not exist. In 2.3b consider the two orders $ABCDBE$ and $ABDCBE$. Here, the tangent is continuous for both orders, but the curvature is only continuous for the order $ABDCBE$. This example shows that the geometric properties may be different for what appears to be the same point set. Something else must be the defining factor.

Theorem 1 (Serret-Frenet formulas). *Let $\boldsymbol{\alpha} : (a,b) \to \mathbb{R}^3$ be a unit-speed curve in $\mathbb{R}^3$ with $k(s) \neq 0$ for all $s \in (a,b)$. Then the following system of equations hold:*

$$\begin{aligned}\mathbf{t}' &= k\mathbf{n}\\ \mathbf{n}' &= -k\mathbf{t} + \tau\mathbf{b}\\ \mathbf{b}' &= -\tau\mathbf{n}\end{aligned} \tag{2.1.4}$$

Proof. The first and last relations are already shown, but it remains to find $\mathbf{n}'$ expressed in terms of the basis $\{\mathbf{t}, \mathbf{n}, \mathbf{b}\}$. This triple form a right-handed orthonormal basis for $\mathbb{R}^3$ and so we have

$$\mathbf{t} \times \mathbf{n} = \mathbf{b}, \qquad \mathbf{n} \times \mathbf{b} = \mathbf{t}, \text{ and } \mathbf{b} \times \mathbf{t} = \mathbf{n}$$

from which we find, using the anticommutative property of the cross product and the product rule of differentiation,

$$\mathbf{n}' = (\mathbf{b} \times \mathbf{t})' = \mathbf{b}' \times \mathbf{t} + \mathbf{b} \times \mathbf{t}' = -\tau(\mathbf{n} \times \mathbf{t}) + k(\mathbf{b} \times \mathbf{n}) = \tau\mathbf{b} - k\mathbf{t}. \qquad \square$$

The equations (2.1.4) are called the Serret-Frenet equations and the orthogonal system $\{\mathbf{t}, \mathbf{n}, \mathbf{b}\}$ the Frenet frame.

We now show in two parts what is known as the *Fundamental Theorem of Space Curves*. Its essence is that a space curve is determined up to a Euclidean motion of $\mathbb{R}^3$ by its curvature and torsion.

Definition 7 (Euclidean motion). Let $\mathbf{y} \in \mathbb{R}^3$. A Euclidean motion of $\mathbb{R}^3$ is an affine map $F : \mathbb{R}^3 \to \mathbb{R}^3$ of the form

$$F(\mathbf{x}) = A\mathbf{x} + \mathbf{y}$$

for all $\mathbf{x} \in \mathbb{R}^3$ where A is an orthogonal transformation.

Theorem 2 (Fundamental theorem, uniqueness). *Let $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ be two unit-speed curves in $\mathbb{R}^3$ defined on the same interval (a,b). Assume that they have the same curvature $k(s) > 0$ and torsion $\tau(s)$ for all $s \in (a,b)$. Then, there is a Euclidean motion F that maps $\boldsymbol{\alpha}_1$ onto $\boldsymbol{\alpha}_2$.*

Proof. Let $\{\mathbf{t}_i, \mathbf{n}_i, \mathbf{b}_i\}$ be the Frenet frame of $\boldsymbol{\alpha}_i$. These are well-defined since we have non-zero curvature. Let s_0 be some fixed value of s. Since

$\{\mathbf{t}_i(s_0), \mathbf{n}_i(s_0), \mathbf{b}_i(s_0)\}$ $(i = 1, 2)$ are two right-handed orthonormal bases for $\mathbb{R}^3$ there is naturally a rotation A that takes $\mathbf{t}_1(s_0), \mathbf{n}_1(s_0)$ and $\mathbf{b}_1(s_0)$ to $\mathbf{t}_2(s_0), \mathbf{n}_2(s_0)$ and $\mathbf{b}_2(s_0)$, respectively. There is also a translation that takes $\boldsymbol{\alpha}_1(s_0)$ to $\boldsymbol{\alpha}_2(s_0)$. It is therefore no problem finding a Euclidean motion F such that

$$\begin{gathered} F(\boldsymbol{\alpha}_1(s_0)) = \boldsymbol{\alpha}_2(s_0), \\ A(\mathbf{t}_1(s_0)) = \mathbf{t}_2(s_0),\ A(\mathbf{n}_1(s_0)) = \mathbf{n}_2(s_0),\ A(\mathbf{b}_1(s_0)) = \mathbf{b}_2(s_0). \end{gathered} \tag{2.1.5}$$

Consider now the real valued function

$$f(s) = \|(A \circ \mathbf{t}_1)(s) - \mathbf{t}_2(s)\|^2 + \|(A \circ \mathbf{n}_1)(s) - \mathbf{n}_2(s)\|^2 + \|(A \circ \mathbf{b}_1)(s) - \mathbf{b}_2(s)\|^2$$

for $s \in (a, b)$. By construction $f(s_0) = 0$ and so if we can show that $f'(s) = 0$ for all s we can conclude that f vanishes, i.e. (2.1.5) holds, for all s. Let us compute $f'(s)$:

$$\begin{aligned} f'(s) = &\quad 2\left((A \circ \mathbf{t}_1)'(s) - \mathbf{t}_2'(s)\right) \cdot \left((A \circ \mathbf{t}_1)(s) - \mathbf{t}_2(s)\right) \\ &+ 2\left((A \circ \mathbf{n}_1)'(s) - \mathbf{n}_2'(s)\right) \cdot \left((A \circ \mathbf{n}_1)(s) - \mathbf{n}_2(s)\right) \\ &+ 2\left((A \circ \mathbf{b}_1)'(s) - \mathbf{b}_2'(s)\right) \cdot \left((A \circ \mathbf{b}_1)(s) - \mathbf{b}_2(s)\right) \\ = &- 2\left((A \circ \mathbf{t}_1)'(s) \cdot \mathbf{t}_2(s) + (A \circ \mathbf{t}_1)(s) \cdot \mathbf{t}_2'(s)\right) \\ &- 2\left((A \circ \mathbf{n}_1)'(s) \cdot \mathbf{n}_2(s) + (A \circ \mathbf{n}_1)(s) \cdot \mathbf{n}_2'(s)\right) \\ &- 2\left((A \circ \mathbf{b}_1)'(s) \cdot \mathbf{b}_2(s) + (A \circ \mathbf{b}_1)(s) \cdot \mathbf{b}_2'(s)\right). \end{aligned}$$

We can now use that $(A \circ \mathbf{t}_1)'(s) = (A \circ \mathbf{t}_1')(s)$ and similarly for $\mathbf{n}_i$ and $\mathbf{b}_i$ in combination with the Serret-Frenet formulas (2.1.4) to find (omitting (s) for readability)

$$\begin{aligned} f'(s) = &- 2\,[k_1 A\mathbf{n}_1 \cdot \mathbf{t}_2 + A\mathbf{t}_1 \cdot \mathbf{n}_2 \\ &+ A(-k_1\mathbf{t}_1 + \tau_1\mathbf{b}_1) \cdot \mathbf{n}_2 + A\mathbf{n}_1 \cdot (-k_2\mathbf{t}_2 + \tau_2\mathbf{b}_2) \\ &+ A(-\tau_1\mathbf{n}_1) \cdot \mathbf{b}_2 + A\mathbf{b}_1 \cdot (-\tau_2\mathbf{n}_2)]\,. \end{aligned}$$

By assumption $k_1(s) = k_2(s)$ and $\tau_1(s) = \tau_2(s)$ and so the expression in the brackets cancels to 0. Hence $(F \circ \boldsymbol{\alpha}_1)'(s) = (A \circ \mathbf{t}_1)(s) = \boldsymbol{\alpha}_2'(s)$ for all s, and so there must exist some $\mathbf{y} \in \mathbb{R}^3$ such that

$$(F \circ \boldsymbol{\alpha}_1)(s) = \boldsymbol{\alpha}_2(s) + \mathbf{y}$$

for all s. By the construction of F we have $\mathbf{y} = \mathbf{0}$, and so the Euclidean motion F maps $\boldsymbol{\alpha}_1$ onto $\boldsymbol{\alpha}_2$. □

Theorem 3 (Fundamental theorem, existence). *Let $k : (a, b) \to \mathbb{R}$ and $\tau : (a, b) \to \mathbb{R}$ be differential functions, and $k > 0$ on the whole domain. Then, there exists a unit-speed curve $\boldsymbol{\alpha} : (a, b) \to \mathbb{R}^3$ whose curvature is k and torsion is τ. For some $s_0 \in (a, b)$ the value $\boldsymbol{\alpha}(s_0)$ can be prescribed arbitrarily, and so can $\mathbf{t}(s_0)$ and $\mathbf{n}(s_0)$ as long as they are of unit length and orthogonal.*

Proof. Let us set up a linear system with initial conditions so that they match the Serret-Frenet formulas (2.1.4). Consider therefore the following system of ODEs

$$\begin{aligned}
\boldsymbol{\alpha}' &= \mathbf{t} \\
\mathbf{t}' &= k\mathbf{n} \\
\mathbf{n}' &= -k\mathbf{t} + \tau\mathbf{b} \\
\mathbf{b}' &= -\tau\mathbf{n}
\end{aligned}$$

which correspond to 12 differential equations, together with some prescribed initial condition

$$\begin{aligned}
\boldsymbol{\alpha}(s_0) &= \boldsymbol{\alpha}_0 = (\alpha_1, \alpha_2, \alpha_3)_0 \\
\mathbf{t}(s_0) &= \mathbf{t}_0 = (t_1, t_2, t_3)_0 \\
\mathbf{n}(s_0) &= \mathbf{n}_0 = (n_1, n_2, n_3)_0 \\
\mathbf{b}(s_0) &= \mathbf{b}_0 = \mathbf{t}_0 \times \mathbf{n}_0 = (t_2 n_3 - t_3 n_2, t_3 n_1 - t_1 n_3, t_1 n_2 - t_2 n_1)_0
\end{aligned}$$

where we require that $\|\mathbf{t}_0\| = \|\mathbf{n}_0\| = 1$ and $\mathbf{t}_0 \cdot \mathbf{n}_0 = 0$. The subscript of zero indicates that the values are constant. The theory of ODEs[2] shows us that such a system has a solution. Hence, the functions k and τ together with the prescribed initial conditions define a unit-speed curve with k and τ as curvature and torsion. □

2.2 Smooth Surfaces

The goal of this section is to precisely define what we mean by a surface. Surfaces are the two-dimensional analogues of curves. Intuitively, a surface should resemble a deformed plane when considering any sufficiently small piece of it. Globally, however, surfaces may be more complicated. See for instance the two examples at the end of this chapter (Figure 2.6 and 2.7).

As touched upon in the introduction we can either study surfaces extrinsically or intrinsically. The extrinsic properties relate to the embedding in Euclidean space, while the intrinsic properties belong to the geometric object itself and will not change depending on how we represent it. We will consider both as there are some fascinating and surprising results as to what we can deduce with only intrinsic reasoning. Most of the reasoning will be carried out by doing measurements along curves on the surface. We begin with some definitions.

Definition 8 (Homeomorphism). The map $f : A \to B$ is called a homeomorphism if it is continuous and bijective and that its inverse map $f^{-1} : B \to A$ is also continuous. If there is a homeomorphism between two spaces, they are said to be homeomorphic.

[2]See for instance Appendix A in *Nonlinear Ordinary Differential Equations* [11] or Appendix I.1 in *An Introduction to Differential Geometry* [12].

Definition 9 (Surface). We call $M \subset \mathbb{R}^3$ a surface if for every point $\mathbf{x} \in M$ there is an open set $D \subset \mathbb{R}^2$ and an open set $W \subset \mathbb{R}^3$ containing $\mathbf{x}$ such that D and $M \cap W$ are homeomorphic.

With this definition, a surface M is equipped with a collection of homeomorphisms

$$\mathbf{X} : D \to M \cap W,$$

which we will call *surface patches*.[3] A collection of surface patches covering the surface is, quite fittingly, called the *atlas* of M. To avoid that different atlases could define the same surface; one can consider the maximal atlas, i.e. the collection of *all* allowable surface patches. Such distinctions are, however, not the most important for the applications here, and we instead show the difference between a surface and a surface patch with a classic example.

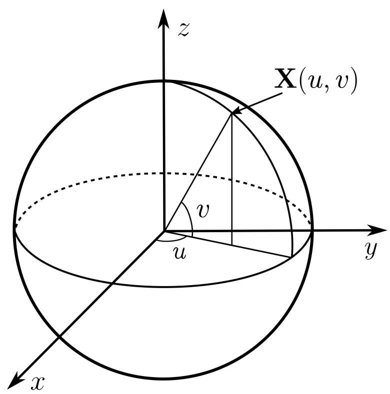


Figure 2.4: The standard patch $\mathbf{X}$ for a sphere with longitude u and latitude v.

Example 2 (The sphere). The sphere of radius $R > 0$ is a surface. With coordinates as in Figure 2.4 the natural parameterisation would be

$$\mathbf{X}(u, v) = (R\cos u \cos v, R\sin u \cos v, R\sin v).$$

To cover the whole sphere one could take $u \in [-\pi, \pi]$ and $v \in [-\pi/2, \pi/2]$, but this does not give an open subset of $\mathbb{R}^2$ to fulfil our definition of a surface patch. Taking the largest open set possible

$$D' = \left\{ (u, v) \in \mathbb{R}^2 | -\pi < u < \pi \text{ and } -\frac{\pi}{2} < v < \frac{\pi}{2} \right\}$$

leaves half a great circle running down the back of Figure 2.4. This means that $\mathbf{X} : D' \to \mathbb{R}^3$ only cover a patch of the sphere. We can define a new

[3] The terms *chart* and *local surface* are also in use.

patch by rotating the first one to cover the missing great circle. These two together will then form an atlas for the sphere. ♣

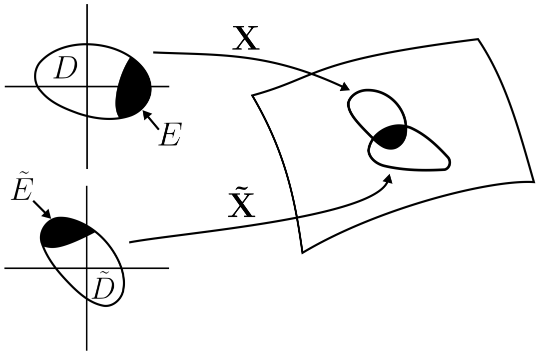


Figure 2.5: Two surface patches $\mathbf{X}$ and $\tilde{\mathbf{X}}$ with overlapping images.

As Example 2 shows, a point on a surface will generally lie in the image of more than one surface patch. Let $\mathbf{X} : D \to M \cap W$ and $\tilde{\mathbf{X}} : \tilde{D} \to M \cap \tilde{W}$ be as in Figure 2.5, and let $\mathbf{x} \in M \cap W \cap \tilde{W}$. Taking the inverse mapping of the overlapping region, $\mathbf{X}^{-1}(M \cap W \cap \tilde{W})$, gives an open set $E \subset D$. Similarly for $\tilde{\mathbf{X}}^{-1}$ gives an open set $\tilde{E} \subset \tilde{D}$. We can now define the *transition map* from $\mathbf{X}$ to $\tilde{\mathbf{X}}$ to be the composite mapping $\mathbf{X}^{-1} \circ \tilde{\mathbf{X}} : \tilde{E} \to E$. If we denote the map by Φ then for all $(\tilde{u}, \tilde{v}) \in \tilde{E}$ we have $\tilde{\mathbf{X}}(\tilde{u}, \tilde{v}) = \mathbf{X}\left(\Phi(\tilde{u}, \tilde{v})\right)$.

We require some structure of the surface to be able to discuss patches and curves on the surface. Let D be an open subset of $\mathbb{R}^m$. We say that the map $f : D \to \mathbb{R}^n$ is of class C^r if all the partial derivatives of order r (including mixed) of all components exist and are continuous. The map is said to be smooth (C^∞) if each component has continuous partial derivatives of all orders. Our focus is the case $m = 2$ and $n = 3$, i.e. $\mathbb{R}^2 \to \mathbb{R}^3$, and to highlight the notation in use we write

$$\mathbf{X}(u, v) = (x(u, v), y(u, v), z(u, v)),$$

$$\frac{\partial \mathbf{X}}{\partial u}(u, v) = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right) = \mathbf{X}_u.$$

Definition 10 (Regular surface patch). A surface patch $\mathbf{X} : D \to \mathbb{R}^3$ is regular if it is a smooth map and $\mathbf{X}_u$ and $\mathbf{X}_v$ are linearly independent for all $(u, v) \in D$.

Remark. This condition is equivalent to $\mathbf{X}_u \times \mathbf{X}_v \neq 0$ for all $(u, v) \in D$ and that the Jacobian matrix of the map has rank 2 for all $(u, v) \in D$.

We are now in a position to define the surfaces that we will study in this thesis. From now on the term ‘surface’ will be used meaning smooth surface.

Definition 11 (Smooth surface). A smooth surface is a surface whose atlas consists of regular surface patches.

2.3 The Tangent Space

To study our surfaces we look at smooth curves that live on the surface. As for space curves we can imagine the path of a moving point, but now forcing the point to stay on the surface. If $\boldsymbol{\alpha} : (a, b) \to \mathbb{R}^3$ is contained in the image of the surface patch $\mathbf{X} : D \to \mathbb{R}^3$, then there is a function $(a, b) \to D$ given by $t \mapsto (u(t), v(t))$ such that

$$\boldsymbol{\alpha}(t) = \mathbf{X}(u(t), v(t)). \tag{2.3.1}$$

So, if $\mathbf{X}$ is a surface patch of a surface, $\boldsymbol{\alpha}$ given by (2.3.1) is a curve on that surface.

With curves on a surface we are able to define tangent vectors and the tangent space.

Definition 12 (Tangent vector). Let M be a surface. A vector $\mathbf{v}$ is called a tangent vector to M at P if there exists a curve on M which passes through P with velocity vector $\mathbf{v}$ at P.

Definition 13 (Tangent space). The tangent space of a surface M at P denoted $T_P M$ is the set of all tangent vectors to M at P.

Remark. This definition, via velocities of curves, is the most intuitive definition of the tangent space and is sufficient for our use. Other definitions, via derivations or the cotangent space exist, but find their uses in a more abstract setting.[4]

Proposition 1. *Let $\mathbf{X} : D \to \mathbb{R}^3$ be a surface patch of the surface M and let P be a point on the image of the patch. Then, the tangent space to the surface M at P is the subspace of $\mathbb{R}^3$ spanned by $\mathbf{X}_u$ and $\mathbf{X}_v$. Here, (u, v) are coordinates in D and the derivatives are evaluated at the point $(u_0, v_0) \in D$ such that $\mathbf{X}(u_0, v_0) = P$*

Proof. For a smooth curve of the form (2.3.1) we have by the chain rule

$$\frac{d}{dt}\boldsymbol{\alpha} = \boldsymbol{\alpha}' = \mathbf{X}_u u' + \mathbf{X}_v v', \tag{2.3.2}$$

[4] See section 2.3 in *Modern Differential Geometry for Physicists* [13] for a full discussion.

and since $\boldsymbol{\alpha}'$ is tangent to the curve it is tangent to the surface. So any tangent vector at the point $P = \mathbf{X}(u_0, v_0)$ is a linear combination of $\mathbf{X}_u(u_0, v_0)$ and $\mathbf{X}_v(u_0, v_0)$.

Conversely, any vector in the subspace of $\mathbb{R}^3$ spanned by $\mathbf{X}_u$ and $\mathbf{X}_v$ has the form $c_1\mathbf{X}_u + c_2\mathbf{X}_v$ for some real coefficients c_1 and c_2. Since D is open, we can always choose t such that $(u_0 + c_1 t, v_0 + c_2 t) \in D$ and hence the curve

$$\boldsymbol{\alpha}(t) = \mathbf{X}(u_0 + c_1 t, v_0 + c_2 t)$$

is smooth on M, and for $t = 0$ we get $\boldsymbol{\alpha}' = a\mathbf{X}_u + b\mathbf{X}_v$. This shows that any vector in the span of $\mathbf{X}_u(u_0, v_0)$ and $\mathbf{X}_v(u_0, v_0)$ is the tangent vector of some some curve on M passing through P. □

2.4 The Surface Normal and Orientability

The 'orientation' of a space curve (2.1.1) is straight forward; it goes from $\boldsymbol{\alpha}(a)$ to $\boldsymbol{\alpha}(b)$, i.e. at each point in the direction of the unit tangent $\mathbf{t}$. For surfaces, we are left with a choice. Thinking of a surface, e.g. the sheet of paper you are looking at, it is easy to believe that it has two sides. For the paper the sides are 'up' and 'down', and for a sphere it is 'in' and 'out'. Such surfaces are said to be orientable, but a surface does not have to be orientable. We now formalise this.

Since we consider regular surfaces, $\mathbf{X}_u$ and $\mathbf{X}_v$ are linearly independent for the surface patch $\mathbf{X}$ and form a plane subspace of $\mathbb{R}^3$. This gives us the opportunity to define a normal to this plane at a point P in the orthogonal complement to the tangent space. Choosing

$$\mathbf{N}_{\mathbf{X}}(u, v) = \frac{\mathbf{X}_u \times \mathbf{X}_v}{\|\mathbf{X}_u \times \mathbf{X}_v\|}(u, v) \tag{2.4.1}$$

and evaluating at the point $P = \mathbf{X}(u_0, v_0)$ gives a unit normal to the surface M at P for the patch $\mathbf{X}$. There is, of course, another unit vector that is orthogonal to both $\mathbf{X}_u$ and $\mathbf{X}_v$, i.e. the vector of opposite sign of $\mathbf{N}_{\mathbf{X}}$.

Proposition 2. *The unit normal $\mathbf{N}_{\mathbf{X}}$ at a point P depends on the choice of surface patch $\mathbf{X}$ covering the point.*

Proof. Let $\mathbf{X} : D \to M \cap W$ and $\tilde{\mathbf{X}} : \tilde{D} \to M \cap \tilde{W}$ be two surface patches for the surface M covering the point $P \in M \cap W \cap \tilde{W}$. Let Φ denote the transition map from $\mathbf{X}$ to $\tilde{\mathbf{X}}$, so that $(u, v) = \Phi(\tilde{u}, \tilde{v})$ where $(u, v) \in D$ and $(\tilde{u}, \tilde{v}) \in \tilde{D}$. The mapping Φ has Jacobian matrix

$$J(\Phi) = \begin{bmatrix} \frac{\partial u}{\partial \tilde{u}} & \frac{\partial u}{\partial \tilde{v}} \\ \frac{\partial v}{\partial \tilde{u}} & \frac{\partial v}{\partial \tilde{v}} \end{bmatrix}.$$

We want to express $\mathbf{N}_{\tilde{\mathbf{X}}}$ in terms of $\mathbf{N}_{\mathbf{X}}$ and therefore look at $\tilde{\mathbf{X}}_{\tilde{u}} \times \tilde{\mathbf{X}}_{\tilde{v}}$. The chain rule gives

$$\begin{aligned}\tilde{\mathbf{X}}_{\tilde{u}} &= \frac{\partial u}{\partial \tilde{u}}\mathbf{X}_u + \frac{\partial v}{\partial \tilde{u}}\mathbf{X}_v,\\ \tilde{\mathbf{X}}_{\tilde{v}} &= \frac{\partial u}{\partial \tilde{v}}\mathbf{X}_u + \frac{\partial v}{\partial \tilde{v}}\mathbf{X}_v.\end{aligned}$$

Using that the cross product is anticommutative and distributive over addition it follows that

$$\begin{aligned}\tilde{\mathbf{X}}_{\tilde{u}} \times \tilde{\mathbf{X}}_{\tilde{u}} =& \left(\frac{\partial u}{\partial \tilde{u}}\mathbf{X}_u + \frac{\partial v}{\partial \tilde{u}}\mathbf{X}_v\right) \times \left(\frac{\partial u}{\partial \tilde{v}}\mathbf{X}_u + \frac{\partial v}{\partial \tilde{v}}\mathbf{X}_v\right)\\ =&\frac{\partial u}{\partial \tilde{u}}\frac{\partial v}{\partial \tilde{v}}\left(\mathbf{X}_u \times \mathbf{X}_v\right) + \frac{\partial v}{\partial \tilde{u}}\frac{\partial u}{\partial \tilde{v}}\left(\mathbf{X}_v \times \mathbf{X}_u\right)\\ =&\left(\frac{\partial u}{\partial \tilde{u}}\frac{\partial v}{\partial \tilde{v}} - \frac{\partial v}{\partial \tilde{u}}\frac{\partial u}{\partial \tilde{v}}\right)\left(\mathbf{X}_u \times \mathbf{X}_v\right)\\ =&\det(J(\Phi))\left(\mathbf{X}_u \times \mathbf{X}_v\right).\end{aligned}$$

From (2.4.1) we find

$$\mathbf{N}_{\tilde{\mathbf{X}}} = \frac{\tilde{\mathbf{X}}_{\tilde{u}} \times \tilde{\mathbf{X}}_{\tilde{u}}}{\|\tilde{\mathbf{X}}_{\tilde{u}} \times \tilde{\mathbf{X}}_{\tilde{u}}\|} = \frac{\det(J(\Phi))\left(\mathbf{X}_u \times \mathbf{X}_v\right)}{|\det(J(\Phi))|\,\|\mathbf{X}_u \times \mathbf{X}_v\|} = \operatorname{sign}\Big(\det\big(J(\Phi)\big)\Big)\mathbf{N}_{\mathbf{X}}.$$

This shows that two different surface patches may have different normal vectors, depending on the sign of the Jacobian of the transition map. □

We now define the term orientable surface and give an example.

Definition 14 (Orientable surface). Let M be a smooth surface and let Φ be any transition map between two patches in the atlas of M. The surface is orientable if $\det(J(\Phi)) > 0$ where Φ is defined.

Example 3 (Möbius strip). The Möbius strip is the standard example of a non-orientable surface, as it is clear that it only has one 'side'. Following a path around the surface leaves you at the same point, but with a surface normal in the opposite direction. Figure 2.6 shows a Möbius strip parameterised by the patch $\mathbf{X}(u,v) = (x(u,v), y(u,v), z(u,v))$ for $(u,v) \in (0, 2\pi) \times (-1, 1)$, where

$$\begin{aligned} x(u,v) &= \frac{v}{2} \sin\left(\frac{u}{2}\right), \\ y(u,v) &= \left(1 + \frac{v}{2}\cos\left(\frac{u}{2}\right)\right)\cos u, \\ z(u,v) &= \left(1 + \frac{v}{2}\cos\left(\frac{u}{2}\right)\right)\sin u. \end{aligned} \tag{2.4.2}$$

♣

Remark. From now on we write $\mathbf{N}_{\mathbf{X}}$ as just $\mathbf{N}$ since there is usually no confusion.

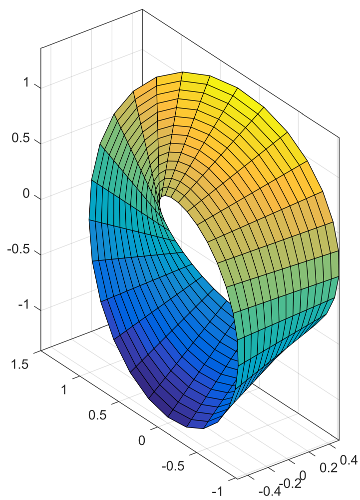


Figure 2.6: The non-orientable Möbius strip from Example 3. The figure is generated in Matlab [7] by the parameterisation (2.4.2).

To conclude this chapter we include another surface plot. It is of an orientable surface that looks like a seashell.

Example 4 (Seashell surface). The surface is given for $(u,v) \in (0,6\pi) \times (0,2\pi)$ by the patch $\mathbf{X}(u,v) = (x,y,z)$, where

$$\begin{aligned} x(u,v) &= 2\left[1 - e^{u/(6\pi)}\right]\cos u \cos^2(v/2), \\ y(u,v) &= 2\left[e^{u/(6\pi)} - 1\right]\sin u \cos^2(v/2), \\ z(u,v) &= 1 - e^{u/(3\pi)} - \sin v + e^{u/(6\pi)}\sin v. \end{aligned} \tag{2.4.3}$$

♣

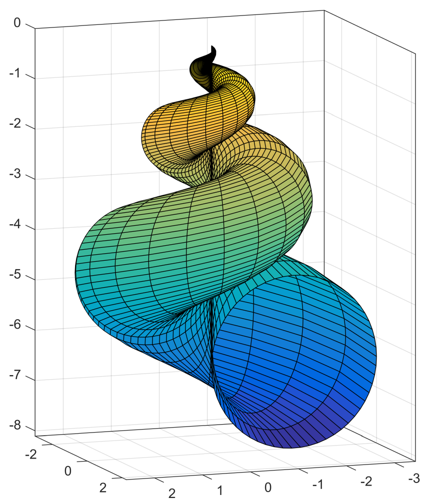


Figure 2.7: The beautiful seashell surface from Example 4. The figure is generated in Matlab [7] by the parameterisation (2.4.3).

2.5 The First Fundamental Form

Now that we have a notion of a surface, it is time to start doing measurements. In Euclidean space, measuring distances is easy: We use the Pythagorean Theorem to find the distance between two points. This approach does, however, not work if you want to measure distances *on* a surface. Using a measuring stick to measure the length between two points on a football seems like a bad strategy. We need something that curves along the surface.

Let us therefore consider the curve $\boldsymbol{\alpha}$ on the surface patch $\mathbf{X}$. From the definition of arc length we have that $ds/dt = \|\boldsymbol{\alpha}'\|$ which we can combine with (2.3.2) to get

$$\begin{aligned}\left(\frac{ds}{dt}\right)^2 &= \|\boldsymbol{\alpha}'\|^2 \\ &= (\mathbf{X}_u u' + \mathbf{X}_v v') \cdot (\mathbf{X}_u u' + \mathbf{X}_v v') \\ &= (\mathbf{X}_u \cdot \mathbf{X}_u)(u')^2 + 2(\mathbf{X}_u \cdot \mathbf{X}_v)u'v' + (\mathbf{X}_v \cdot \mathbf{X}_v)(v')^2 \\ &= E\left(\frac{du}{dt}\right)^2 + 2F\left(\frac{du}{dt}\frac{dv}{dt}\right) + G\left(\frac{dv}{dt}\right)^2.\end{aligned}$$

Here we have introduced, in the notation of Gauss[5],

$$E = \mathbf{X}_u \cdot \mathbf{X}_u, \quad F = \mathbf{X}_u \cdot \mathbf{X}_v \text{ and } G = \mathbf{X}_v \cdot \mathbf{X}_v. \tag{2.5.1}$$

With this in place, we can find the length, L, of a curve $\boldsymbol{\alpha} : (a, b) \to \mathbb{R}^3$ where $\boldsymbol{\alpha}(t) = \mathbf{X}(u(t), v(t))$ by

$$L = \int_a^b \sqrt{E\left(\frac{du}{dt}\right)^2 + 2F\left(\frac{du}{dt}\frac{dv}{dt}\right) + G\left(\frac{dv}{dt}\right)^2}\, dt.$$

Definition 15 (First fundamental form). The first fundamental form is the quadratic from given by

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

where E, F and G are functions given by (2.5.1).

The first fundamental form can be thought of as the generalisation of the infinitesimal Pythagorean Theorem $ds^2 = dx^2 + dy^2$ [14]. Also, taking the square root of ds^2 gives the infinitesimal arc length ds. Let us now show two well-known metrics.

Example 5 (Polar coordinates). It is well known that the polar coordinates r, θ for the Euclidean plane relates to the Cartesian coordinates x, y by the formulas

$$x = r\cos\theta, \text{ and } y = r\sin\theta.$$

Consider therefore the mapping

$$\mathbf{X}(r, \theta) = (r\cos\theta, r\sin\theta, 0).$$

This gives

$$\begin{aligned}\mathbf{X}_r &= (\cos\theta, \sin\theta, 0), \\ \mathbf{X}_\theta &= (-r\sin\theta, r\cos\theta, 0)\end{aligned}$$

[5] See for example Gauss's own paper 'Disquisitiones generales circa superficies curvas' from 1827 [4].

and for the coefficients of the first fundamental form:

$$\begin{aligned}
E &= \mathbf{X}_r \cdot \mathbf{X}_r = \cos^2\theta + \sin^2\theta = 1,\\
F &= \mathbf{X}_r \cdot \mathbf{X}_\theta = -r\cos\theta\sin\theta + r\cos\theta\sin\theta = 0,\\
G &= \mathbf{X}_\theta \cdot \mathbf{X}_\theta = r^2\cos^2\theta + r^2\sin^2\theta = r^2.
\end{aligned}$$

We thus see that the first fundamental form is the well known formula

$$ds^2 = dr^2 + r^2 d\theta^2.$$

♣

Example 6 (The sphere). For R a positive constant, consider as in Example 2 the patch

$$\mathbf{X}(u,v) = (R\cos u\cos v, R\sin u\cos v, R\sin v),$$

where we take $u \in (-\pi,\pi)$ and $v \in (-\pi/2,\pi/2)$. We want to compute the first fundamental form, and so with the formulas

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

we compute

$$\begin{aligned}
E = \mathbf{X}_u \cdot \mathbf{X}_u &= R^2\sin^2 u\cos^2 v + R^2\cos^2 u\cos^2 v = R^2\cos^2 v\\
F = \mathbf{X}_u \cdot \mathbf{X}_v &= R^2\sin u\sin v\cos u\cos v - R^2\sin u\sin v\cos u\cos v = 0\\
E = \mathbf{X}_v \cdot \mathbf{X}_v &= R^2\cos^2 u\sin^2 v + R^2\sin^2 u\sin^2 v + R^2\cos^2 v\\
&= R^2\left(\sin^2 v(\cos^2 u + \sin^2 u) + \cos^2 v\right) = R^2
\end{aligned}$$

and find the metric for the sphere to be

$$ds^2 = R^2\cos^2 v du^2 + R^2 dv^2.$$

♣

Let us look at what happens to a small area element, dA, on a given patch $\mathbf{X}$. The area spanned by the two infinitesimal vectors $\mathbf{X}_u du$ and $\mathbf{X}_v dv$ are given by the cross product $dA = \|\mathbf{X}_u du \times \mathbf{X}_v dv\| = \|\mathbf{X}_u \times \mathbf{X}_v\| dudv$. This suggests the following definition.

Definition 16 (Surface area). Let $\mathbf{X} : D \to \mathbb{R}^3$ be a smooth surface patch and let R be a compact subset of $\mathbf{X}(D)$. The area $A_{\mathbf{X}}(R)$ of the region R is then given by

$$A_{\mathbf{X}}(R) = \iint_{\mathbf{X}^{-1}(R)} \|\mathbf{X}_u \times \mathbf{X}_v\| dudv. \tag{2.5.2}$$

Remark. The expression $dudv$ in (2.5.2) has a different meaning than that of the first fundamental form. While the $dudv$ in the first fundamental form comes from the symmetric product $\mathbf{X}_u \cdot \mathbf{X}_v$, the one for dA is determined by an antisymmetric product and should be written $du \times dv$ or with the wedge (exterior product) $du \wedge dv$.

Theorem 4. *The first fundamental form enables us to measure lengths, angles, and areas in a surface.*

Proof. For lengths and angles, it is sufficient to show that the first fundamental form determines an inner product on the tangent space. Let us model it on the dot product of $\mathbb{R}^3$. Let $\mathbf{v} = a\mathbf{X}_u + b\mathbf{X}_v$ and $\mathbf{w} = c\mathbf{X}_u + d\mathbf{X}_v$ be two arbitrary tangent vectors at a point P of the surface M. Here a, b, c, d are real numbers. The dot product of two such vectors are:

$$\begin{aligned}
\mathbf{v} \cdot \mathbf{w} &= (a\mathbf{X}_u + b\mathbf{X}_v) \cdot (c\mathbf{X}_u + d\mathbf{X}_v) \\
&= Eac + F(ad + bc) + Gbd \\
&= \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} E & F \\ F & G \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}.
\end{aligned}$$

Since we, for a given point and patch, can represent the vectors $\mathbf{v}$ and $\mathbf{w}$ by just $v = (a, b)$ and $w = (c, d)$ we can readily define the inner product

$$\langle v, w \rangle_P = v^T \begin{bmatrix} E & F \\ F & G \end{bmatrix} w.$$

This shows that the matrix $\left[\begin{smallmatrix} E & F \\ F & G \end{smallmatrix}\right]$ determines an inner product, and hence lengths and angles, of tangent vectors.

For the area we prove that

$$\|\mathbf{X}_u \times \mathbf{X}_v\| = (EG - F^2)^{1/2}, \tag{2.5.3}$$

as this shows that the area in (2.5.2) is given by the coefficients of the first fundamental form. From vector calculus we know

$$\begin{aligned}
\|\mathbf{X}_u \times \mathbf{X}_v\| &= \|\mathbf{X}_u\| \|\mathbf{X}_v\| \sin\theta, \\
\mathbf{X}_u \cdot \mathbf{X}_v &= \|\mathbf{X}_u\| \|\mathbf{X}_v\| \cos\theta,
\end{aligned}$$

so adding the squares of these identities gives

$$\|\mathbf{X}_u \times \mathbf{X}_v\|^2 + (\mathbf{X}_u \cdot \mathbf{X}_v)^2 = \|\mathbf{X}_u\|^2 \|\mathbf{X}_v\|^2 (\sin^2\theta + \cos^2\theta) = \|\mathbf{X}_u\|^2 \|\mathbf{X}_v\|^2.$$

Rearranging shows the result

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

Since the area of a surface patch is unchanged under a reparameterisation[6], it is clear that one can divide a surface, M, into pieces that each are contained in a single surface patch, then use (2.5.2) to calculate the area of each piece and finally add each contribution to find the total area of M. □

[6] We refer to Proposition 5.3 in [8] for a proof.

Chapter 3

Curvature of Surfaces

In this section, we explore the curvature of a surface from an extrinsic viewpoint, before concluding that one type of curvature is intrinsic. Describing the curvature of a space curve was intuitive. By Lemma 1 the curvature at a point measure of how much a curve differs from being a straight line at that point. For a surface, we measure how much a small patch fails at being a plane. We also use curves confined to the surface to define the Gaussian and mean curvature.

3.1 Tensor Notation and the Einstein Summation Convention

Before continuing the discussion, we change the notation slightly. The reasons for changing notation is to save space and also have a more robust notation that easily extends to a higher-dimensional setting. Instead of using local coordinates u and v we instead write u^1 and u^2, and also define

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

so that $E = g_{11}$, $F = g_{12} = g_{21}$ and $G = g_{22}$. Further, the determinant of the matrix with coefficients g_{ij} will appear in some formulas and is given the symbol

$$g = \det(g_{ij}).$$

One advantage of this formulation is that we can adopt the Einstein summation convention and omit the summation symbol in any sum where the index of summation appears as both a subscript and a superscript. In effect, we write

$$g_{ij}du^i du^j \text{ instead of } \sum_{i,j} g_{ij}du^i du^j.$$

A downside could have been writing the derivative as $u^{j\prime}$, but we combat this using dot-notation, $\dot{u}^j$ when possible.

The inner product of two tangent vectors $\mathbf{v} = v^i\mathbf{X}_i$ and $\mathbf{w} = w^i\mathbf{X}_i$ at P on the surface M, i.e. $\mathbf{v}, \mathbf{w} \in T_PM$, is now

$$\langle v, w\rangle_P = \mathbf{v} \cdot \mathbf{w} = \sum_{i,j} v^i w^j \mathbf{X}_i \cdot \mathbf{X}_j = g_{ij}v^i w^j.$$

We also introduce the notation g^{ij} for the element i, j of the inverse of the matrix (g_{ij})

$$(g^{ij}) = (g_{ij})^{-1}$$

so that for the two-dimensional case

$$(g^{ij}) = \frac{1}{g}\begin{bmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{bmatrix}.$$

3.2 The Second Fundamental Form

Let us now take an external view and try to relate the curvature of a curve on a surface with the geometry of the surface itself. Consider the smooth curve $\boldsymbol{\alpha}(s) = \mathbf{X}(u(s), v(s))$, where s is arc length. We defined the curvature of a space curve to be the length of $\boldsymbol{\alpha}''$ and therefore decompose this vector into two parts,

$$\boldsymbol{\alpha}'' = \boldsymbol{\alpha}''_{\text{tan}} + \boldsymbol{\alpha}''_{\text{nor}},$$

one tangent and one normal to the surface. For this section, denote d/ds with a dot $\cdot$ to simplify notation.

With the summation convention in place, the chain rule gives

$$\dot{\boldsymbol{\alpha}} = \dot{u}^i \mathbf{X}_i$$

and

$$\begin{aligned} \frac{d}{ds}\dot{\boldsymbol{\alpha}} = \ddot{\boldsymbol{\alpha}} &= \ddot{u}^i \mathbf{X}_i + \dot{u}^i \frac{d}{ds}\mathbf{X}_i \\ &= \ddot{u}^k \mathbf{X}_k + \dot{u}^i \dot{u}^j \mathbf{X}_{ij}. \end{aligned} \tag{3.2.1}$$

The acceleration tangent to the surface will be spanned by the tangent vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ while the acceleration normal to the surface will be a scalar of the unit normal $\mathbf{N}$ given by (2.4.1). Hence, we split $\mathbf{X}_{ij}$ into components with these as basis vectors and define functions Γ^k_{ij} and b_{ij} by the *Gauss formulas*

$$\mathbf{X}_{ij} = \Gamma^k_{ij}\mathbf{X}_k + b_{ij}\mathbf{N}. \tag{3.2.2}$$

Substituting (3.2.2) into (3.2.1) gives

$$\ddot{\boldsymbol{\alpha}} = (\ddot{u}^k + \Gamma^k_{ij}\dot{u}^i \dot{v}^j)\mathbf{X}_k + (b_{ij}\dot{u}^i \dot{u}^j)\mathbf{N} \tag{3.2.3}$$

and so

$$\ddot{\boldsymbol{\alpha}}_{\text{tan}} = (\ddot{u}^k + \Gamma^k_{ij}\dot{u}^i \dot{v}^j)\mathbf{X}_k, \tag{3.2.4}$$

$$\ddot{\boldsymbol{\alpha}}_{\text{nor}} = (b_{ij}\dot{u}^i \dot{u}^j)\mathbf{N}. \tag{3.2.5}$$

Notice that taking the dot product of $\mathbf{X}_{ij}$, defined in (3.2.2), and the surface normal, $\mathbf{N}$, gives

$$b_{ij} = \mathbf{X}_{ij} \cdot \mathbf{N} \tag{3.2.6}$$

since $\mathbf{N}$ is orthogonal to $\mathbf{X}_1$ and $\mathbf{X}_2$. The values of b_{ij} are independent of the curve under consideration and the quadratic form $b_{ij}du^i du^j$ is called the second fundamental form.

Definition 17 (The second fundamental form). The second fundamental form of a surface patch is the quadratic form given by

$$b_{11}(du^1)^2 + 2b_{12}du^1du^2 + b_{22}(du^2)^2,$$

where b_{ij} is given by (3.2.6).

Let us show a geometric interpretation of this formula. To do so, consider first the unit-speed curve $\boldsymbol{\alpha}$ of the parameter s. When s increases to $s+\Delta s$ the curve moves off the tangent at $\boldsymbol{\alpha}(s)$ by $(\boldsymbol{\alpha}(s+\Delta s) - \boldsymbol{\alpha}(s)) \cdot \mathbf{n}$ where $\mathbf{n}$ is the principle normal vector. For sufficiently small Δs the quadratic approximation

$$\boldsymbol{\alpha}(s+\Delta s) \approx \boldsymbol{\alpha}(s) + \dot{\boldsymbol{\alpha}}(s)\Delta s + \frac{1}{2}\ddot{\boldsymbol{\alpha}}(s)(\Delta s)^2$$

is decent by Taylor's theorem. Now, since $\dot{\boldsymbol{\alpha}} \cdot \mathbf{n} = 0$ and $\ddot{\boldsymbol{\alpha}} \cdot \mathbf{n} = k\mathbf{n}\cdot\mathbf{n} = k$ for unit-speed curves, we find

$$(\boldsymbol{\alpha}(s+\Delta s) - \boldsymbol{\alpha}(s)) \cdot \mathbf{n} \approx \frac{1}{2}k(\Delta s)^2.$$

This shows that the deviation from the tangent line of $\boldsymbol{\alpha}$ is dominated by the term $k(\Delta s)^2$.

Moving over to the surface patch $\mathbf{X}$ with unit normal $\mathbf{N}$, let (u^1, u^2) be the parameters of the patch. As they change slightly to $(u^1+\Delta u^1, u^2+\Delta u^2)$ the surface moves from its tangent plane at $\mathbf{X}(u^1, u^2)$ by

$$\left(\mathbf{X}(u^1+\Delta u^1, u^2+\Delta u^2) - \mathbf{X}(u^1, u^2)\right) \cdot \mathbf{N}.$$

The idea is shown in Figure 3.1.

The same approximation, but now in two variables, gives

$$\begin{aligned}&\mathbf{X}(u^1+\Delta u^1, u^2+\Delta u^2) - \mathbf{X}(u^1, u^2) \approx \\ &\quad \mathbf{X}_1\Delta u^1 + \mathbf{X}_2\Delta u^2 + \frac{1}{2}\left(\mathbf{X}_{11}(\Delta u^1)^2 \ +2\mathbf{X}_{12}\Delta u^1\Delta u^2 + \mathbf{X}_{22}(\Delta u^2)^2\right).\end{aligned}$$

Taking the scalar product with the unit normal, and using that $\mathbf{X}_i \cdot \mathbf{N} = 0$ and $\mathbf{X}_{ij} \cdot \mathbf{N} = b_{ij}$ results in

$$\begin{aligned}&\left(\mathbf{X}(u^1-\Delta u^1, u^2+\Delta u^2) - \mathbf{X}(u^1, u^2)\right) \cdot \mathbf{N} \approx \\ &\frac{1}{2}\left(b_{11}(\Delta u^1)^2 + 2b_{12}\Delta u^1\Delta u^2 + b_{22}(\Delta u^2)^2\right).\end{aligned}$$

This shows that the second fundamental form is in some sense to a surface patch what curvature is to a unit-speed curve.

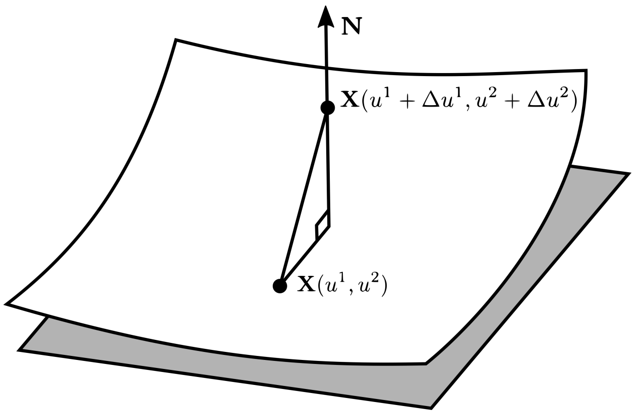


Figure 3.1: A slight variation of the initial point $\mathbf{X}(u^1, u^2)$ moves the surface patch away from the tangent plane. The tangent plane at $\mathbf{X}(u^1, u^2)$ is illustrated in grey.

3.3 Principal, Mean and Gaussian Curvature

Let us continue the local study of the surface patch $\mathbf{X}$. Let it be in the atlas of the surface M, and consider a fixed point P on the surface. At P, the surface normal together with any tangent vector $\mathbf{v} \in T_P M$ define a plane. The intersection of this plane and the surface gives a new curve $\boldsymbol{\alpha}_{\mathbf{v}}$ called the normal section of M at P in the direction $\mathbf{v}$. See Figure 3.2 for an illustration of the idea.

The curvature of the normal section is called the normal curvature, for which we take the following formal definition:

Definition 18 (Normal curvature). Let $\mathbf{v} = v^i \mathbf{X}_i$ be a tangent vector to the surface M at P, i.e. $\mathbf{v} \in T_P M$. Then, the normal curvature of M at P in the direction $\mathbf{v}$ is given by

$$k_n(\mathbf{v}) = \frac{b_{ij} v^i v^j}{g_{mn} v^m v^n}. \tag{3.3.1}$$

Proposition 3. *Let $\boldsymbol{\alpha}(s) = \mathbf{X}(u^1(s), u^2(s))$ be a unit-speed curve on M with surface patch $\mathbf{X}$. Further, let the curve satisfy $\boldsymbol{\alpha}(s_0) = P$ and $\dot{\boldsymbol{\alpha}}(s_0) = \mathbf{v}$ where $\mathbf{v} \in T_P M$ then*

$$k_n(\mathbf{v}) = \ddot{\boldsymbol{\alpha}} \cdot \mathbf{N}.$$

Proof. First, $\dot{\boldsymbol{\alpha}}(s_0) = \dot{u}^i \mathbf{X}_i(u^1(s_0), u^2(s_0))$ so that $v^i = \dot{u}^i$. Second, since $\boldsymbol{\alpha}(s)$ is unit-speed, $1 = \|\dot{\boldsymbol{\alpha}}(s_0)\|^2 = \|\mathbf{v}\|^2 = g_{mn} v^m v^n$, equation (3.3.1) gives

$$k_n(\mathbf{v}) = b_{ij} \dot{u}^i \dot{u}^j.$$

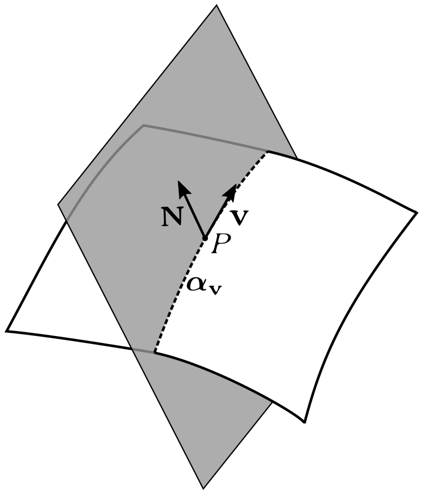


Figure 3.2: A surface patch with a normal section $\boldsymbol{\alpha}_{\mathbf{v}}$ (dotted line) defined by the intersection of the surface with the plane (grey) generated by the surface normal $\mathbf{N}$ and tangent vector $\mathbf{v}$ at P.

That $b_{ij}\dot{u}^i\dot{u}^j = \ddot{\boldsymbol{\alpha}} \cdot \mathbf{N}$ follows directly from (3.2.3), since

$$\ddot{\boldsymbol{\alpha}} \cdot \mathbf{N} = (\ddot{\boldsymbol{\alpha}}_{\tan} + \ddot{\boldsymbol{\alpha}}_{\text{nor}}) \cdot \mathbf{N} = \ddot{\boldsymbol{\alpha}}_{\text{nor}} \cdot \mathbf{N} = b_{ij}\dot{u}^i\dot{u}^j. \qquad \square$$

Varying $\mathbf{v}$ gives different normal sections $\boldsymbol{\alpha}_{\mathbf{v}}$ and hence different normal curvatures. The maximum and minimum value of the normal curvature are important and given individual names.

Definition 19 (Principal curvatures). Let $k_1 \geq k_2$ be the maximum and minimum values of the normal curvatures, $k_n(\mathbf{v})$, of the surface M at P. Then, k_1 and k_2 are called the principle curvatures of M at P. The directions for which we find these, are called the principle directions.

Definition 20 (Mean curvature). The average of the principle curvatures,

$$H = H(P) = \frac{1}{2}(k_1 + k_2),$$

is the mean curvature of M at P.

Definition 21 (Gaussian curvature). The product of the principle curvatures,

$$K = K(P) = k_1 k_2,$$

is the Gaussian curvature of M at P.

Lemma 2. *The Gaussian curvature is given by*

$$K = k_1 k_2 = \frac{b}{g},$$

where $b = b_{11}b_{22} - b_{12}^2$ *and* $g = g_{11}g_{22} - g_{12}^2$.

Proof. The proof is based on Lagrange multipliers. We want to find the extreme values of the normal curvature, $k_n(\mathbf{v})$, and consider only unit-speed curves giving the constraint $g_{mn}v^m v^n = 1$. We therefore construct the Lagrangian

$$\mathcal{L}(v^1, v^2, \lambda) = \frac{b_{ij}v^i v^j}{g_{mn}v^m v^n} - \lambda(g_{mn}v^m v^n - 1).$$

Where the Lagrangian $\mathcal{L}$ is stationary we have $\nabla_{v^1,v^2,\lambda}\mathcal{L} = \mathbf{0}$. The partial derivative with respect to λ is just the constraint $g_{mn}v^m v^n = 1$ which in turn simplify the equations. We get

$$\frac{1}{2}\frac{\partial \mathcal{L}}{\partial v^1} = b_{11}v^1 + b_{12}v^2 - \lambda g_{11}v^1 - \lambda g_{12}v^2 = 0 \tag{3.3.2}$$

and

$$\frac{1}{2}\frac{\partial \mathcal{L}}{\partial v^2} = b_{12}v^1 + b_{22}v^2 - \lambda g_{12}v^1 - \lambda g_{22}v^2 = 0. \tag{3.3.3}$$

Taking v^1 times (3.3.2) and adding v^2 times (3.3.3) results in

$$\begin{aligned} b_{11}(v^1)^2 + b_{12}v^1v^2 - \lambda g_{11}(v^1)^2 - \lambda g_{12}v^1v^2 \\ +b_{12}v^1v^2 + b_{22}(v^2)^2 - \lambda g_{12}v^1v^2 - \lambda g_{22}(v^2)^2 = 0. \end{aligned}$$

Rearranging this gives that $\lambda = k_n(\mathbf{v})$ and so we have the following system to solve, writing just k_n for the normal curvature:

$$\begin{bmatrix} b_{11} - k_n g_{11} & b_{12} - k_n g_{12} \\ b_{12} - k_n g_{12} & b_{22} - k_n g_{22} \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{3.3.4}$$

A non-zero solution to this require the determinant to be zero, and so doing the multiplication reveals

$$k_n^2(g_{11}g_{22} - (g_{12})^2) - k_n(g_{12}b_{12} + g_{22}b_{11} - g_{11}b_{22}) + b_{11}b_{22} - (b_{12})^2 = 0.$$

With the definition of $g = g_{11}g_{22} - (g_{12})^2$ and $b = b_{11}b_{22} - (b_{12})^2$ in place, this equals

$$k_n^2 - k_n\frac{g_{12}b_{12} + g_{22}b_{11} - g_{11}b_{22}}{g} + \frac{b}{g} = 0.$$

Now, since the principal curvatures, k_1 and k_2, are the roots of this polynomial they must satisfy $(k_n - k_1)(k_n - k_2) = 0$ showing that

$$k_1 k_2 = \frac{b}{g}.$$

□

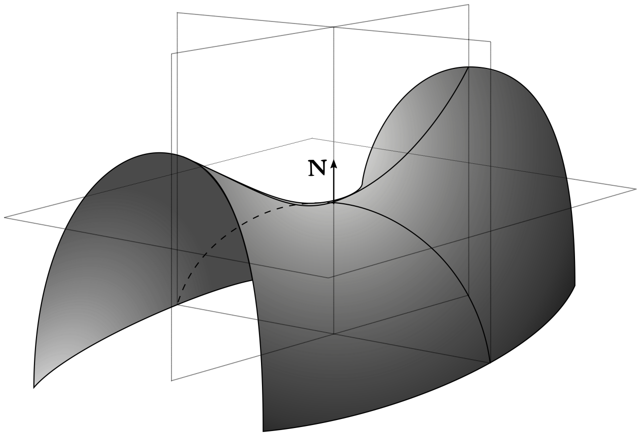


Figure 3.3: A saddle-like surface (hyperbolic paraboloid) for which the principal curvatures have opposite sign, and the Gaussian curvature is negative.

Example 7 (Gaussian curvature of some surfaces)**.** For a plane any normal section is a straight line. This means that the normal curvature is everywhere zero, and so also the Gaussian curvature. For the sphere of radius $R > 0$ any normal section is a great circle with curvature $1/R$, and so the Gaussian curvature of a sphere is $1/R^2$. For the hyperbolic paraboloid in Figure 3.3 the curves of principal curvatures at the origin are shown. Notice how these two curves curve in opposite direction, giving a negative Gaussian curvature. ♣

We now present a remarkable[1] result by Gauss. It is, according to Alfred Gray, 'one of the most celebrated theorems of the 19th century' [14]. Its essence is that the Gaussian curvature of a surface is an intrinsic property, independent of the embedding in $\mathbb{R}^3$. This is in no way obvious, as each of the normal curvatures are defined extrinsically.

Theorem 5 (Gauss's Theorema Egregium)**.** *The Gaussian curvature is an intrinsic quantity only dependent on the first fundamental form.*

Remark. Gauss's original proof found in [4] is not the simplest and is omitted. We do, however, include a result which is directly applicable to the discussion of the geodesics, and serves as an alternative proof.

Corollary 1. *The Gaussian curvature of a (smooth) surface is given by the formula*

$$K = -\frac{1}{\sqrt{g}}\left[\frac{\partial}{\partial u^1}\left(\frac{\sqrt{g}}{g_{11}}\Gamma^2_{12}\right) - \frac{\partial}{\partial u^2}\left(\frac{\sqrt{g}}{g_{11}}\Gamma^2_{11}\right)\right], \tag{3.3.5}$$

[1]The name 'remarkable' has stuck after it was called 'Theorema Egregium' (remarkable theorem) by Gauss himself in his *Disquisitiones generales circa superficies curvas* [4]

where the Γ^k_{ij}'s are the Christoffel symbols defined by the Gauss formulas (3.2.2).

Proof. The proof is based on chapter 46 in *Differential Geometry* by Erwin Kreyszig [10]. We want to find a different expression for

$$K = \frac{b}{g}$$

and start by rewriting b. Differentiating the relation $\mathbf{X}_i \cdot \mathbf{N} = 0$ gives

$$\mathbf{X}_{ij} \cdot \mathbf{N} + \mathbf{X}_i \cdot \frac{\partial \mathbf{N}}{\partial u^j} = 0$$

so that

$$b_{ij} = -\mathbf{X}_i \cdot \mathbf{N}_j,$$

when using (3.2.6) and defining $\mathbf{N}_j = \partial \mathbf{N}/\partial u^j$. This gives

$$\begin{aligned} b = b_{11}b_{22} - b_{12}^2 &= (-\mathbf{X}_1 \cdot \mathbf{N}_1)(-\mathbf{X}_2 \cdot \mathbf{N}_2) - (\mathbf{X}_1 \cdot \mathbf{N}_2)(\mathbf{X}_2 \cdot \mathbf{N}_1) \\ &= (\mathbf{X}_1 \times \mathbf{X}_2) \cdot (\mathbf{N}_1 \times \mathbf{N}_2), \end{aligned}$$

and hence, by using (2.4.1) and (2.5.3), we find

$$\sqrt{g}K = \frac{b}{\sqrt{g}} = \mathbf{N} \cdot (\mathbf{N}_1 \times \mathbf{N}_2).$$

Let us now introduce an arbitrary unit vector, $\mathbf{T}$, in the tangent space of the surface and define the vector

$$\mathbf{M} = \mathbf{T} \times \mathbf{N} \text{ so that } \mathbf{N} = \mathbf{M} \times \mathbf{T}.$$

This gives

$$\sqrt{g}K = (\mathbf{M} \times \mathbf{T}) \cdot (\mathbf{N}_1 \times \mathbf{N}_2) = (\mathbf{M} \cdot \mathbf{N}_1)(\mathbf{T} \cdot \mathbf{N}_2) - (\mathbf{M} \cdot \mathbf{N}_2)(\mathbf{T} \cdot \mathbf{N}_1). \quad (3.3.6)$$

Differentiating the relations $\mathbf{M} \cdot \mathbf{N} = 0$, $\mathbf{T} \cdot \mathbf{N} = 0$ and $\mathbf{M} \cdot \mathbf{M} = 1$ shows that

$$\mathbf{M} \cdot \mathbf{N}_i = -\mathbf{M}_i \cdot \mathbf{N}, \quad \mathbf{T} \cdot \mathbf{N}_i = -\mathbf{T}_i \cdot \mathbf{N}, \quad \mathbf{M} \cdot \mathbf{M}_i = 0.$$

This helps us rewrite (3.3.6) in the following way:

$$\begin{aligned} \sqrt{g}K &= (\mathbf{M}_1 \cdot \mathbf{N})(\mathbf{T}_2 \cdot \mathbf{N}) - (\mathbf{M}_2 \cdot \mathbf{N})(\mathbf{T}_1 \cdot \mathbf{N}) \\ &= \mathbf{M}_1 \cdot [(\mathbf{T}_2 \cdot \mathbf{N})\mathbf{N}] - \mathbf{M}_2 \cdot [(\mathbf{T}_1 \cdot \mathbf{N}) \cdot \mathbf{N}] \\ &= \mathbf{M}_1 \cdot [(\mathbf{T}_2 \cdot \mathbf{N})\mathbf{N} + (\mathbf{T}_2 \cdot \mathbf{M})\mathbf{M}] - \mathbf{M}_2 \cdot [(\mathbf{T}_1 \cdot \mathbf{N}) \cdot \mathbf{N} + (\mathbf{T}_1 \cdot \mathbf{M}) \cdot \mathbf{M}] \\ &= \mathbf{M}_1 \cdot \mathbf{T}_2 - \mathbf{M}_2 \cdot \mathbf{T}_1 \\ &= \mathbf{M}_1 \cdot \mathbf{T}_2 + \mathbf{M} \cdot \mathbf{T}_{12} - \mathbf{M}_2 \cdot \mathbf{T}_1 - \mathbf{M} \cdot \mathbf{T}_{12} \\ &= \frac{\partial}{\partial u^1}(\mathbf{M} \cdot \mathbf{T}_2) - \frac{\partial}{\partial u^2}(\mathbf{M} \cdot \mathbf{T}_1), \end{aligned}$$

so that

$$K = \frac{1}{\sqrt{g}} \left(\frac{\partial}{\partial u^1}(\mathbf{M} \cdot \mathbf{T}_2) - \frac{\partial}{\partial u^2}(\mathbf{M} \cdot \mathbf{T}_1) \right). \tag{3.3.7}$$

We are free to choose the unit vector in the tangent space, $\mathbf{T}$, but choosing

$$\mathbf{T} = \frac{\mathbf{X}_1}{\|\mathbf{X}_1\|} = \frac{\mathbf{X}_1}{\sqrt{g_{11}}}$$

gives

$$\mathbf{T}_i = \frac{\mathbf{X}_{1i}}{\sqrt{g_{11}}} + \mathbf{X}_1 \frac{\partial}{\partial u^i}\left(\frac{1}{\sqrt{g_{11}}}\right)$$

so that

$$\mathbf{M} \cdot \mathbf{T}_i = \left(\frac{\mathbf{X}_1}{\sqrt{g_{11}}} \times \mathbf{N}\right) \cdot \left(\frac{\mathbf{X}_{1i}}{\sqrt{g_{11}}} + \mathbf{X}_1 \frac{\partial}{\partial u^i}\left(\frac{1}{\sqrt{g_{11}}}\right)\right).$$

Now, since $(\mathbf{X}_1 \times \mathbf{N}) \cdot \mathbf{X}_1 = 0$ we get

$$\begin{aligned} \mathbf{M} \cdot \mathbf{T}_i &= \left(\mathbf{N} \times \frac{\mathbf{X}_1}{\sqrt{g_{11}}}\right) \cdot \frac{\mathbf{X}_{1i}}{\sqrt{g_{11}}} \\ &= \frac{1}{g_{11}}(\mathbf{X}_{1i} \times \mathbf{X}_1) \cdot \mathbf{N}, \end{aligned} \tag{3.3.8}$$

where in the last equality we used the cyclic property of the vector triple product.

Recall the Gauss formulas (3.2.2):

$$\mathbf{X}_{ij} = \Gamma^k_{ij}\mathbf{X}_k + b_{ij}\mathbf{N}.$$

Taking the vector product of this formula with $\mathbf{X}_1$ gives

$$\begin{aligned} \mathbf{X}_{1i} \times \mathbf{X}_1 &= \Gamma^1_{1i}(\mathbf{X}_1 \times \mathbf{X}_1) + \Gamma^2_{1i}(\mathbf{X}_2 \times \mathbf{X}_1) + b_{1i}(\mathbf{N} \times \mathbf{X}_1) \\ &= -\Gamma^2_{1i}\sqrt{g}\mathbf{N}, \end{aligned}$$

and then combining with (3.3.8) shows that

$$\mathbf{M} \cdot \mathbf{T}_i = -\frac{\sqrt{g}}{g_{11}}\Gamma^2_{1i}.$$

Insert this into (3.3.7) to arrive at

$$K = \frac{1}{\sqrt{g}} \left[\frac{\partial}{\partial u^1}\left(-\frac{\sqrt{g}}{g_{11}}\Gamma^2_{12}\right) - \frac{\partial}{\partial u^2}\left(-\frac{\sqrt{g}}{g_{11}}\Gamma^2_{11}\right) \right],$$

which is the result after pulling the minus sign outside the brackets. □

Remark. In the next chapter, we show that the Christoffel symbols only depend on the coefficients of the first fundamental form and their derivatives. This means that they are intrinsic, and hence, this derivation provides an alternative proof of the Theorema Egregium.

Remark. We can find a similar expression to (3.3.5), involving g_{22} and Γ^1_{ij} instead of g_{11} and Γ^2_{ij}, by a slight modification of the derivation above.

Chapter 4

Geodesics

As indicated earlier (see for example (3.2.3)-(3.2.5)), the curvature of a curve $\boldsymbol{\alpha}$ on a surface M comes from two sources. First, if the surface itself is curved then $\boldsymbol{\alpha}$ generally bends along M as seen from the surrounding 3-dimensional space. An easy example to visualise is a bug living on a sphere going straight ahead. Albeit going 'straight' it will follow a great circle, and so its path will curve since the sphere is curved. And second, regardless of the curvature of M, the curve $\boldsymbol{\alpha}$ may bend within M. This latter curvature is called the geodesic curvature. An example here is a circle on a flat plane. This curve has curvature that does not come from the surface itself, as the plane is flat.

In this chapter, we begin by showing that the Christoffel symbols only depend on the first fundamental form. We then turn to discussing the geodesic curvature and the geodesic equations and show the connection between vanishing geodesic curvature and the length-minimising curves on a surface. After demonstrating that there are geodesics in any direction radiating from a point on a surface, we use this to introduce coordinate systems on the surface where the coordinate curves are geodesics. These new coordinates help us prove the Gauss-Bonnet theorem. The chapter ends giving a strategy to measure the Gaussian curvature for the two-dimensional being mentioned in the introduction.

4.1 The Christoffel Symbols

Before going further with the discussion of the geodesics curve, we take a pause to discuss the Christoffel symbols[1]. In particular, Γ^r_{ij} defined by the Gauss formulas (3.2.2). It turns out that the Christoffel symbols belong to the intrinsic geometry of M despite being defined extrinsically.

Theorem 6 (Intrinsic property of Christoffel symbols). *The Christoffel symbols Γ^r_{ij} depend only on the coefficients of the first fundamental form.*

Proof. To help with the calculations define the Christoffel symbols of the first kind by

$$\Gamma_{ijk} = \Gamma^r_{ij} g_{rk},$$

[1]What we call the Christoffel symbols here are more precisely the Christoffel symbols of the second kind. They were originally introduced by Elwin Bruno Christoffel in 'Über die Transformation der homogenen Differentialausdrücke zweiten Grades' [15]

so that $\Gamma_{ij}^r = \Gamma_{ijk} g^{rk}$. Notice that this is symmetric in its first two indices. Take now the Gauss formulas

$$\mathbf{X}_{ij} = \Gamma_{ij}^r \mathbf{X}_r + b_{ij} \mathbf{N}_{\mathbf{X}}$$

and take the scalar product with $\mathbf{X}_k$

$$\mathbf{X}_{ij} \cdot \mathbf{X}_k = \Gamma_{ij}^r \mathbf{X}_r \cdot \mathbf{X}_k = \Gamma_{ij}^r g_{rk} = \Gamma_{ijk}. \tag{4.1.1}$$

Taking the partial derivative of g_{ik} with respect to u^j gives

$$\frac{\partial g_{ik}}{\partial u^j} = \frac{\partial}{\partial u^j}(\mathbf{X}_i \cdot \mathbf{X}_k) = \mathbf{X}_{ij} \cdot \mathbf{X}_k + \mathbf{X}_i \cdot \mathbf{X}_{kj} = \Gamma_{ijk} + \Gamma_{kji},$$

and changing the indices gives the similar relations

$$\frac{\partial g_{ji}}{\partial u^k} = \Gamma_{jki} + \Gamma_{ikj} = \Gamma_{kji} + \Gamma_{ikj}$$

and

$$\frac{\partial g_{kj}}{\partial u^i} = \Gamma_{kij} + \Gamma_{jik} = \Gamma_{ikj} + \Gamma_{ijk}.$$

Collecting in a clever way shows that

$$\begin{aligned}\frac{\partial g_{ik}}{\partial u^j} + \frac{\partial g_{kj}}{\partial u^i} - \frac{\partial g_{ji}}{\partial u^k} &= (\Gamma_{ijk} + \Gamma_{kji}) + (\Gamma_{ikj} + \Gamma_{ijk}) - (\Gamma_{kji} + \Gamma_{ikj}) \\ &= 2\Gamma_{ijk},\end{aligned}$$

or explicitly written

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial u^j} + \frac{\partial g_{kj}}{\partial u^i} - \frac{\partial g_{ji}}{\partial u^k}\right).$$

To get back to Γ_{ij}^r, multiply by g^{kr} and sum over k:

$$\Gamma_{ij}^r = \frac{1}{2} g^{kr}\left(\frac{\partial g_{ik}}{\partial u^j} + \frac{\partial g_{jk}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^k}\right). \tag{4.1.2}$$

This shows that the Christoffel symbols are given in terms of the g_{ij}'s and their derivatives. □

4.2 Geodesic Curvature and the Geodesic Equations

Let us return to the discussion of the curve $\boldsymbol{\alpha}(s) = \mathbf{X}(u^1(s), u^2(s))$ on the surface M. As before, let s be arc length and a dot denote d/ds. Since we defined

$$\ddot{\boldsymbol{\alpha}} = \ddot{\boldsymbol{\alpha}}_{\text{tan}} + \ddot{\boldsymbol{\alpha}}_{\text{nor}} = (\ddot{u}^r + \Gamma_{ij}^r \dot{u}^i \dot{u}^j)\mathbf{X}_r + (b_{ij}\dot{u}^i \dot{u}^j)\mathbf{N}$$

it follows that $\ddot{\boldsymbol{\alpha}}_{\tan}$ is orthogonal to both $\mathbf{N}$ and $\dot{\boldsymbol{\alpha}}$. Clearly $\ddot{\boldsymbol{\alpha}}_{\tan} \cdot \mathbf{N} = 0$ and $\ddot{\boldsymbol{\alpha}}_{\tan} \cdot \dot{\boldsymbol{\alpha}} = (\ddot{\boldsymbol{\alpha}}_{\tan} + \ddot{\boldsymbol{\alpha}}_{\text{nor}}) \cdot \dot{\boldsymbol{\alpha}} = \ddot{\boldsymbol{\alpha}} \cdot \dot{\boldsymbol{\alpha}} = \frac{1}{2}\frac{d}{ds}(\dot{\boldsymbol{\alpha}} \cdot \dot{\boldsymbol{\alpha}}) = \frac{1}{2}\frac{d}{ds}1 = 0$ where we use that $\dot{\boldsymbol{\alpha}}$ is of unit length. Therefore, $\ddot{\boldsymbol{\alpha}}_{\tan}$ is proportional to the unit vector $\mathbf{N} \times \dot{\boldsymbol{\alpha}}$. Let us call the proportionality factor the geodesic curvature.

Definition 22 (Geodesic curvature). Let $\boldsymbol{\alpha}(s)$ be a unit-speed curve on the surface M with local patch $\mathbf{X}$ and normal vector $\mathbf{N}$. Then, the geodesic curvature is the function k_g defined by

$$\ddot{\boldsymbol{\alpha}}_{\tan} = k_g(\mathbf{N} \times \dot{\boldsymbol{\alpha}}). \tag{4.2.1}$$

There is another way to express this. Take the dot product of (4.2.1) and the vector $\mathbf{N} \times \dot{\boldsymbol{\alpha}}$. This gives $k_g = \ddot{\boldsymbol{\alpha}}_{\tan} \cdot (\mathbf{N} \times \dot{\boldsymbol{\alpha}}) = \ddot{\boldsymbol{\alpha}} \cdot (\mathbf{N} \times \dot{\boldsymbol{\alpha}})$, which is a vector triple product. Such products exhibit cyclic permutations and we can thus write

$$k_g = \mathbf{N} \cdot (\ddot{\boldsymbol{\alpha}} \times \dot{\boldsymbol{\alpha}}). \tag{4.2.2}$$

Theorem 7 (Intrinsic property of the geodesic curvature). *The geodesic curvature k_g, given in (4.2.2), of a curve $\boldsymbol{\alpha}$ contained in the surface patch $\mathbf{X}$ of M depends only on the first fundamental form of M.*

Proof. A direct computation of $\mathbf{N} \cdot (\ddot{\boldsymbol{\alpha}} \times \dot{\boldsymbol{\alpha}})$ shows the result. Recall $\dot{\boldsymbol{\alpha}} = \dot{u}^r \mathbf{X}_r$ and equation (3.2.3) for $\ddot{\boldsymbol{\alpha}}$. This gives

$$\begin{aligned}
\dot{\boldsymbol{\alpha}} \times \ddot{\boldsymbol{\alpha}} =& \dot{u}^r \mathbf{X}_r \times \big((\ddot{u}^k + \Gamma^k_{ij}\dot{u}^i\dot{v}^j)\mathbf{X}_k + (b_{ij}\dot{u}^i\dot{u}^j)\mathbf{N}\big) \\
=& (\dot{u}^1\mathbf{X}_1 + \dot{u}^2\mathbf{X}_2) \times \\
& \big((\ddot{u}^1 + \Gamma^1_{ij}\dot{u}^i\dot{u}^j)\mathbf{X}_1 + (\ddot{u}^2 + \Gamma^2_{ij}\dot{u}^i\dot{u}^j)\mathbf{X}_2 + (b_{ij}\dot{u}^i\dot{u}^j)\mathbf{N}\big)\,.
\end{aligned}$$

Now $\mathbf{X}_i \times \mathbf{X}_i = \mathbf{0}$, and from (2.4.1) we have $\mathbf{X}_1 \times \mathbf{X}_2 = \|\mathbf{X}_1 \times \mathbf{X}_2\|\mathbf{N}$ and in the new notation equation (2.5.3) gives $\|\mathbf{X}_1 \times \mathbf{X}_2\| = \sqrt{g}$. This gives

$$\mathbf{X}_1 \times \mathbf{X}_2 = \sqrt{g}\mathbf{N} \text{ and } \mathbf{X}_2 \times \mathbf{X}_1 = -\sqrt{g}\mathbf{N}.$$

In total $\mathbf{N} \cdot (\ddot{\boldsymbol{\alpha}} \times \dot{\boldsymbol{\alpha}})$ gets the form

$$\mathbf{N} \cdot \big(\dot{u}^1(\ddot{u}^2 + \Gamma^2_{ij}\dot{u}^i\dot{u}^j)\sqrt{g}\mathbf{N} - \dot{u}^2(\ddot{u}^1 + \Gamma^1_{ij}\dot{u}^i\dot{u}^j)\sqrt{g}\mathbf{N} + \dot{u}^r b_{ij}\dot{u}^i\dot{u}^j(\mathbf{X}_r \times \mathbf{N})\big)\,.$$

Use now $\mathbf{N} \cdot \mathbf{N} = 1$ and for the last term $\mathbf{N} \cdot (\mathbf{X}_r \times \mathbf{N}) = \mathbf{X}_r \cdot (\mathbf{N} \times \mathbf{N}) = 0$ to arrive at

$$\begin{aligned}
k_g = \sqrt{g}\,\big[&\Gamma^2_{11}\,(\dot{u}^1)^3 + (2\Gamma^2_{12} - \Gamma^1_{11})(\dot{u}^1)^2\dot{u}^2 \\
&-(2\Gamma^1_{12} - \Gamma^2_{22})\dot{u}^1(\dot{u}^2)^2 - \Gamma^1_{22}(\dot{u}^2)^3 + \dot{u}^1\ddot{u}^2 - \dot{u}^2\ddot{u}^1\big]
\end{aligned} \tag{4.2.3}$$

with everything written out. By Theorem 6, the Christoffel symbols and hence k_g is given by the first fundamental form of the surface M. $\square$

Remark. Although k_g was defined for unit-speed curves it is easy to change the parameter from arc length, s, to some allowable parameter t. Denote by primes derivatives with respect to t then for the last term in the parenthesis:

$$\ddot{u}^1\dot{u}^2 = u^{1\prime\prime}u^{2\prime}(\frac{dt}{ds})^3 = u^{1\prime\prime}u^{2\prime}\dot{t}^3.$$

Similarly, every term in (4.2.3) would get the factor $\dot{t}^3$.

We are now in a position to define the straight curves on a surface: the geodesics. We will call a curve a geodesic if it has no acceleration tangential to the surface.

Definition 23 (Geodesic curve). Let M be a surface and $\boldsymbol{\alpha}$ a unit-speed curve on M. The curve is a geodesic on M if $\ddot{\boldsymbol{\alpha}}_{\tan} = \mathbf{0}$ at every point of $\boldsymbol{\alpha}$.

From the above, $\boldsymbol{\alpha}$ is a geodesic if either

$$\ddot{u}^r + \Gamma^r_{ij}\dot{u}^i\dot{u}^j = 0, \quad r = 1, 2 \tag{4.2.4}$$

or

$$k_g = \mathbf{N} \cdot (\dot{\boldsymbol{\alpha}} \times \ddot{\boldsymbol{\alpha}}) = 0$$

holds. The equations in (4.2.4) are called the Geodesic equations.

4.3 Length-Minimising Properties of Geodesics

In the last section we focused on geodesics as 'straightest' curves; now we show their character as 'shortest'. We consider the problem of finding the curve of shortest length connecting two points on a surface, and conclude that the geodesic equations are the Euler-Lagrange-equations of this minimisation problem.

Theorem 8. *Let $\boldsymbol{\alpha} : [a, b] \to M$ be a curve on the surface M with metric g_{ij}. If $\boldsymbol{\alpha}$ is the shortest curve on M connecting $P = \boldsymbol{\alpha}(a)$ and $Q = \boldsymbol{\alpha}(b)$, then $\boldsymbol{\alpha}$ is a geodesic.*

Proof. The length of $\boldsymbol{\alpha}$ is

$$L(\boldsymbol{\alpha}) = \int_a^b \sqrt{g_{ij}\dot{u}^i\dot{u}^j}\,dt, \tag{4.3.1}$$

which is invariant under a reparameterisation. The assumption that $\boldsymbol{\alpha}$ is the shortest curve connecting P and Q translates to

$$L(\boldsymbol{\alpha}) \leq L(\boldsymbol{\beta}) \text{ for all } \boldsymbol{\beta} : [a, b] \to M \text{ with } \boldsymbol{\beta}(a) = P \text{ and } \boldsymbol{\beta}(b) = Q.$$

Let us compare the length of the curve $\boldsymbol{\alpha}$ with that of its variations, $\boldsymbol{\beta}_\epsilon$, of the form (for $i = 1, 2$)

$$u^i(t) + \epsilon\eta^i(t) \text{ with } \eta^i(a) = \eta^i(b) = 0 \qquad \text{for } i = 1, 2.$$

Here, ϵ is small in absolute value and u^i are the local coordinates. Since for $\epsilon = 0$ we have the shortest curve, we have

$$L(\boldsymbol{\beta}_0) \leq L(\boldsymbol{\beta}_\epsilon) \text{ for all } \epsilon.$$

This in turn gives the condition

$$\frac{d}{d\epsilon} L(\boldsymbol{\beta}_\epsilon)\bigg|_{\epsilon=0} = 0.$$

Let us show that this condition is equivalent to the geodesic equations (4.2.4) by repeating a famous derivation credited to Joseph-Louis Lagrange. It becomes clearer in a general setting, so consider the integral

$$I(\boldsymbol{\alpha}) = \int_a^b f(\dot{\mathbf{u}}(t), \mathbf{u}(t))dt$$

where we write $\mathbf{u} = \mathbf{u}(t) = (u^1(t), u^2(t)))$ to shorten notation. We find

$$0 = \frac{d}{d\epsilon} L(\boldsymbol{\beta}_\epsilon)\bigg|_{\epsilon=0} = \frac{d}{d\epsilon} \int_a^b f(\dot{\mathbf{u}} + \epsilon\dot{\boldsymbol{\eta}}, \mathbf{u} + \epsilon\boldsymbol{\eta})dt\bigg|_{\epsilon=0}$$

where we also write $\boldsymbol{\eta} = (\eta^1(t), \eta^2(t))$ to keep it compact. Bringing the derivative inside the integral and evaluating at $\epsilon = 0$ results in

$$0 = \int_a^b \left(\frac{\partial f}{\partial \dot{u}^i}\dot{\eta}^i + \frac{\partial f}{\partial u^i}\eta^i\right) dt \text{ for } i = 1, 2.$$

Integration by parts and using that $\boldsymbol{\eta}(a) = \boldsymbol{\eta}(b) = \mathbf{0}$ gives the condition

$$0 = \int_a^b \left(-\frac{d}{dt} f_{\dot{u}^i} + f_{u^i}\right) \eta^i dt. \tag{4.3.2}$$

Here, the subscripts indicate which argument of f the partial derivative is taken with resepect to. Since (4.3.2) holds for all $\boldsymbol{\eta}$ it must be true, by the fundamental lemma of calculus of variations (see e.g. lemma 1.1.1 in [16]), that

$$-\frac{d}{dt} f_{\dot{u}^i}(\dot{\mathbf{u}}(t), \mathbf{u}(t)) + f_{u^i}(\dot{\mathbf{u}}(t), \mathbf{u}(t)) = 0$$

and, taking the derivative with respect to t,

$$f_{\dot{u}^i\dot{u}^j}(\dot{\mathbf{u}}(t), \mathbf{u}(t))\ddot{u}^j(t) + f_{\dot{u}^i u^j}(\dot{\mathbf{u}}(t), \mathbf{u}(t))\dot{u}^j(t) = f_{u^i}(\dot{\mathbf{u}}(t), \mathbf{u}(t)) \text{ for } i = 1, 2.$$

Let us insert for

$$f(\dot{\mathbf{u}}(t), \mathbf{u}(t)) = \sqrt{g_{ij}(\mathbf{u}(t))\dot{u}^i(t)\dot{u}^j(t)}$$

to get to to the original problem (4.3.1). We find the system of equations

$$\frac{1}{f}\left(g_{rk}\ddot{u}^r + \frac{\partial g_{ik}}{\partial u^j}\dot{u}^i\dot{u}^j\right) - \frac{1}{f^2}\frac{df}{dt}g_{kr}\dot{u}^r = \frac{1}{2f}\frac{\partial g_{ij}}{\partial u^k}\dot{u}^i\dot{u}^j$$

for $k = 1, 2$. Choosing a curve of constant speed gives f =const. so that $df/dt = 0$ and the equations reduce to

$$\begin{aligned} 0 =& \frac{1}{2}\frac{\partial g_{ij}}{\partial u^k}\dot{u}^i\dot{u}^j - \frac{\partial g_{ik}}{\partial u^j}\dot{u}^i\dot{u}^j - g_{rk}\ddot{u}^r \\ =& \left[\frac{1}{2}\frac{\partial g_{ij}}{\partial u^k} - \frac{\partial g_{ik}}{\partial u^j}\right]\dot{u}^i\dot{u}^j - g_{rk}\ddot{u}^r. \end{aligned}$$

Equation (4.1.1) lets us introduce the Christoffel symbols into this expression. Note that $\Gamma_{ikj}\dot{u}^i\dot{u}^j = \Gamma_{jki}\dot{u}^i\dot{u}^j$ since this is just interchanging summation indices. We get

$$\begin{aligned} 0 &= \left[\frac{1}{2}(\Gamma_{ikj} + \Gamma_{jki}) - (\Gamma_{ijk} + \Gamma_{kji})\right]\dot{u}^i\dot{u}^j - g_{rk}\ddot{u}^r \\ &= \left(\frac{1}{2}\Gamma_{jki} + \frac{1}{2}\Gamma_{jki} - \Gamma_{jki}\right)\dot{u}^i\dot{u}^j - \Gamma_{ijk}\dot{u}^i\dot{u}^j - g_{rk}\ddot{u}^r \end{aligned}$$

so that shortest curves must satisfy

$$g_{rk}\ddot{u}^r + \Gamma_{ijk}\dot{u}^i\dot{u}^j = 0.$$

We can multiply by g^{rk} and sum over k to get

$$\ddot{u}^r + \Gamma^r_{ij}\dot{u}^i\dot{u}^j = 0.$$

This shows that $\boldsymbol{\alpha}$ satisfy the geodesic equations, and thus is a geodesic curve. □

Remark. The converse of this theorem does not always hold. It is in general not true that a geodesic gives the minimum distance between two end points on a surface. Example: The geodesics on a sphere are great circles. For two points, not diametrically opposed, there is obviously two portions of the great circle through the points. One short, and one longer, where the longer takes the long way around the sphere. The longer curve is a geodesic, but it is not the curve of shortest length. When we loosen the restrictions of the smoothness of the patch, there is no guarantee that a geodesic exists between two points. The most common example being the plane with the origin removed and trying to connect two points symmetric around the origin.

We now include a theorem stating that through any point and in any direction on a surface there is a unique geodesic. This allows us to define geodesic coordinate systems in the next section.

Theorem 9. *Let $\mathbf{P}$ be a point on M and $\mathbf{v}$ a unit tangent vector at $\mathbf{P}$. Then, there is a unique geodesic $\boldsymbol{\alpha}$ such that $\boldsymbol{\alpha}(s_0) = \mathbf{P}$ and $\dot{\boldsymbol{\alpha}}(s_0) = \mathbf{v}$.*

Proof. Let us write out the geodesic equations in (4.2.4):

$$\ddot{u}^1 + \Gamma^1_{11}(\dot{u}^1)^2 + 2\Gamma^1_{12}\dot{u}^1\dot{u}^2 + \Gamma^1_{22}(\dot{u}^2)^2 = 0,$$

$$\ddot{u}^2 + \Gamma^2_{11}(\dot{u}^1)^2 + 2\Gamma^2_{12}\dot{u}^1\dot{u}^2 + \Gamma^2_{22}(\dot{u}^2)^2 = 0.$$

These are second order non-linear ODEs of the form

$$\begin{aligned} \ddot{u}^1 &= f_1(u^1, u^2, \dot{u}^1, \dot{u}^2), \\ \ddot{u}^2 &= f_2(u^1, u^2, \dot{u}^1, \dot{u}^2), \end{aligned} \tag{4.3.3}$$

where f_1 and f_2 are smooth functions of the variables, and in particular Lipschitz. So, by the Picard-Lindelöf theorem (see e.g. chapter 8.2 in [17]), there exists a unique solution to (4.3.3) satisfying

$$\begin{aligned} u^1(s_0) &= a, \quad u^2(s_0) = b, \\ \dot{u}^1(s_0) &= c, \quad \dot{u}^2(s_0) = d, \end{aligned} \tag{4.3.4}$$

for constants a, b, c, d and being such that u^i is smooth in a small neighbourhood of s_0.

Suppose that P is in the patch $\mathbf{X}(u^1, u^2)$ of M. Let $P = \mathbf{X}(a, b)$ and $\mathbf{v} = c\mathbf{X}_1(a, b) + d\mathbf{X}_2(a, b)$. The curve $\boldsymbol{\alpha}(s)$ passes through P at $s = s_0$ if and only if $u^1(s_0) = a$ and $u^2(s_0) = b$, and it has tangent vector $\mathbf{v}$ if and only if

$$c\mathbf{X}_1 + d\mathbf{X}_2 = \mathbf{v} = \dot{\boldsymbol{\alpha}}(s_0) = \dot{u}^1(s_0)\mathbf{X}_1 + \dot{u}^2(s_0)\mathbf{X}_2,$$

i.e. $\dot{u}^1(s_0) = c$ and $\dot{u}^2(s_0) = d$. Finding a geodesic passing through P at $s = s_0$ with tangent vector $\mathbf{v}$ at P is, thus, equivalent to solving (4.3.3) with initial values (4.3.4), and since this has a unique solution the proof is complete. □

4.4 Geodesic Coordinates

This section is in preparation for the fascinating Gauss-Bonnet theorem. To prove the theorem, we will use our results for the geodesics to introduce new coordinates. We seek a coordinate system where the coordinate curves, i.e. the curves where only one coordinate varies, are geodesics. We now show that such a coordinate system exists locally around a point and how it simplifies the expressions for the Gaussian curvature and the geodesic curvature. We follow Erwin Kreyzsig's treatment in *Differential Geometry* [10].

Definition 24. A coordinate system on a surface is orthogonal if and only if for any point on the surface

$$g_{12} = 0.$$

Proposition 4. *In orthogonal coordinates the Gaussian curvature takes the form*

$$K = -\frac{1}{\sqrt{g_{11}g_{22}}}\left[\frac{\partial}{\partial u^1}\left(\frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1}\right) + \frac{\partial}{\partial u^2}\left(\frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{11}}}{\partial u^2}\right)\right]. \tag{4.4.1}$$

Proof. For $g_{12} = 0$ it should be clear that $g = \sqrt{g_{11}g_{22}}$ and $g_{ii} = 1/g^{ii}$. This proposition is a special case of (3.3.5), and we therefore compute Γ^2_{11} and Γ^2_{12} from (4.1.2):

$$\Gamma^2_{12} = \frac{1}{2}g^{k2}\left(\frac{\partial g_{1k}}{\partial u^2} + \frac{\partial g_{2k}}{\partial u^1} - \frac{\partial g_{12}}{\partial u^k}\right) = \frac{1}{2g_{22}}\frac{\partial g_{22}}{\partial u^1}$$

and similarly

$$\Gamma^2_{11} = \frac{-1}{2g_{22}}\frac{\partial g_{11}}{\partial u^2}.$$

This gives

$$\frac{\sqrt{g}}{g_{11}}\Gamma^2_{12} = \frac{\sqrt{g_{11}g_{22}}}{g_{11}}\frac{1}{2g_{22}}\left(\frac{\partial g_{22}}{\partial u^1}\right) = \frac{1}{\sqrt{g_{11}}}\frac{1}{2\sqrt{g_{22}}}\frac{\partial g_{22}}{\partial u^1} = \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1}$$

and

$$\frac{\sqrt{g}}{g_{11}}\Gamma^2_{11} = \frac{-1}{\sqrt{g_{22}}}\frac{\partial\sqrt{g_{11}}}{\partial u^2}.$$

Inserting these two relations into (4.1.2) yields the result. □

Let us study the coordinate curves, curves where one of the coordinates are fixed while the other varies. In particular, let $u^1 = \text{const.}$ so that $\dot{u}^1 = 0$ and $\dot{u}^2 = 1/\sqrt{g_{22}}$. This significantly simplifies the expression for the geodesic curvature (4.2.3):

$$k_g = -\Gamma^1_{22}\frac{\sqrt{g}}{g_{22}^{3/2}} \quad \text{when } u^1 = \text{const.} \tag{4.4.2}$$

If we in addition have orthogonal coordinates, Γ^1_{22} has a simple expression. The result is that

$$k_g = \frac{1}{2g_{22}\sqrt{g_{11}}}\frac{\partial g_{22}}{\partial u^1} \quad \text{when } u^1 = \text{const. and } g_{12} = 0. \tag{4.4.3}$$

The same can be done with $u^2 = \text{const.}$ which yields

$$k_g = \Gamma^2_{11}\frac{\sqrt{g}}{g_{11}^{3/2}} \quad \text{when } u^2 = \text{const.} \tag{4.4.4}$$

and

$$k_g = -\frac{1}{2g_{11}\sqrt{g_{22}}}\frac{\partial g_{11}}{\partial u^2} \quad \text{when } u^2 = \text{const. and } g_{12} = 0. \tag{4.4.5}$$

Definition 25 (Field of geodesics). A one-parameter family of geodesics on a surface M is called a field of geodesics in $M' \subset M$ if through every point of M' there passes exactly one of those geodesics.

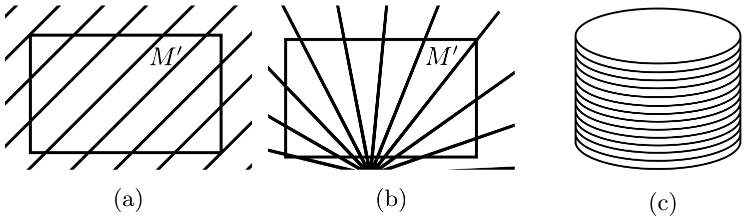


Figure 4.1: (a) and (b) depicts two fields of geodesics in the plane. Recall that geodesics in the plane are straight lines. (a) is a family of parallel lines while (b) is straight lines with origin at a point not contained in M'. (c) depicts the generating lines of a cylinder, which form a field of geodesics on the whole cylinder.

This definition is illustrated in Figure 4.1. We remark that there is no field of geodesics for the whole sphere, since any two great circles intersect. For a smaller patch, we may, of course, find such a field.

For a sufficiently small region of a surface, we may introduce geodesic parallel coordinates in the following way. Choose a field of geodesics as coordinate curves $\hat{u}^2 = \text{const.}$ and take their orthogonal trajectories as the coordinate curves $\hat{u}^1 = \text{const.}$ Since $\hat{u}^1$ and $\hat{u}^2$ are orthogonal, $\hat{g}_{12} = 0$. Also, since it is a geodesic the geodesic curvature vanishes and (4.4.5) gives that $\partial\hat{g}_{11}/\partial\hat{u}^2 = 0$ so that $\hat{g}_{11}$ depends only on $\hat{u}^1$. We can therefore let

$$u^1 = \int_0^{\hat{u}^1} \sqrt{\hat{g}_{11}}d\hat{u}^1, \quad u^2 = \hat{u}^2$$

which is a transformation of coordinates that leaves the geodesic parallel curves unchanged.

This shows that a first fundamental form corresponding to the coordinates u^1, u^2 from above is given by

$$ds^2 = (du^1)^2 + g_{22}(u^1, u^2)(du^2)^2. \tag{4.4.6}$$

We now introduce Riemann normal coordinates in the following proposition.

Proposition 5. *Through an arbitrary point P on a smooth surface M there is exactly one geodesic in every direction. These geodesics form a field of geodesics in $U \setminus P$ where U is a small neighbourhood of P.*

Proof. This proposition is similar to Theorem 9, but let us take a different approach. Introduce on the surface the coordinates $\tilde{u}^1, \tilde{u}^2$ with origin at P satisfying $\tilde{g}_{ij} = \delta_{ij}$. Take $\tilde{u}^i = h^i(s; c^1, c^2)$ $(i = 1, 2)$ to be the solutions to the geodesic equations (4.2.4) satisfying the initial conditions

$$\tilde{u}^1(0) = 0, \qquad \left.\frac{\partial \tilde{u}^1}{\partial s}\right|_{s=0} = c^1$$

$$\tilde{u}^2(0) = 0, \qquad \left.\frac{\partial \tilde{u}^2}{\partial s}\right|_{s=0} = c^2.$$

We want to show that the solutions $h^i(s; c^1, c^2)$ may be represented on the form $\phi^i(sc^1, sc^2)$, for some functions ϕ^i. Consider the parameter transform $s = kt$. This transformation leaves the geodesic equations unchanged and hence if $h^i(s; c^1, c^2)$ is a solution, then so is $h^i(t; \tilde{c}^1, \tilde{c}^2)$ for new constants $\tilde{c}^i$. Now, by the chain rule and boundary conditions

$$\frac{\partial \tilde{u}^i}{\partial t} = \frac{\partial \tilde{u}^i}{\partial s} k \quad \text{and} \quad \left.\frac{\partial \tilde{u}^i}{\partial t}\right|_{t=0} = c^i k \quad \text{for } i = 1, 2$$

showing that $h^i(s; c^1, c^2) = h^i(t; c^1 k, c^2 k)$. For the case $t = 1$ we have $h^i(s; c^1, c^2) = h^i(k; c^1, c^2) = h^i(1; c^1 k, c^2 k)$ and then, inserting s for k gives

$$h^i(s; c^1, c^2) = h^i(1; c^1 s, c^2 s) := \phi^i(v^1, v^2),$$

where we have defined $v^i = c^i s$. Hence, the transformation $\tilde{u}^i = \phi^i(v^1, v^2)$ introduces the coordinates v^1, v^2 on M. These are called Riemann normal coordinates or geodesic normal coordinates on M with center P.

Let us show that this transformation is allowable in a neighbourhood of P. The boundary condition is

$$c^i = \left.\frac{\partial \tilde{u}^i}{\partial s}\right|_{s=0}$$

and by the chain rule

$$\left.\frac{\partial h^i}{\partial s}\right|_{s=0} = \left(\frac{\partial \phi^i}{\partial v^j}\frac{\partial v^j}{\partial s}\right)_{v^1=v^2=0} = \left(\frac{\partial \phi^i}{\partial v^j} c^j\right)_{v^1=v^2=0}.$$

Combining these shows that

$$c^1 = \frac{\partial \phi^1}{\partial v^1} c^1 + \frac{\partial \phi^1}{\partial v^2} c^2$$

$$c^2 = \frac{\partial \phi^2}{\partial v^1} c^1 + \frac{\partial \phi^2}{\partial v^2} c^2$$

where the partial derivatives are evaluated at $v^1 = v^2 = 0$. This gives

$$\left.\frac{\partial \phi^i}{\partial v^j}\right|_{v^1=v^2=0} = \delta_{ij}$$

and hence the Jacobian of the transform at P is the identity matrix. Since the Jacobian is continuous its determinant does not vanish at a sufficiently small neighbourhood of P.

The geodesics through P have the representation $v^i = c^i s$, and assuming that $(c^1)^2 + (c^2)^2 = 1$ the parameter s is the arc length of those geodesics. The values of s, c^1, c^2 corresponding to a point in $U \setminus P$ is obviously uniquely determined. In other words: there exist only one geodesic joining this point and P. This is precisely what it means to have a field of geodesics on $U \setminus P$. $\square$

Returning to the coordinates $\tilde{u}^1, \tilde{u}^2$ in the proof above, we note that these have origin at P and that the first fundamental form satisfy $\tilde{g}_{ij} = \delta_{ij}$ at P. Call the coefficients of the first fundamental form of the coordinates v^1, v^2 by $\hat{g}_{ij}$. At P these must be equal showing that

$$\hat{g}_{ij}|_{v^1=v^2=0} = \delta_{ij}.$$

We now go on to discussing geodesic polar coordinates, which is an important player in the proof of the Gauss-Bonnet theorem. Let v^1, v^2 be Riemann normal coordinates with centre P. If we set

$$v^1 = u^1 \cos u^2, \qquad v^2 = u^1 \sin u^2 \tag{4.4.7}$$

the coordinates u^1, u^2 are called the geodesic polar coordinates on M with centre at P. In this coordinate system, the coordinate curves $u^2 =$ const. correspond to the geodesics through P and the curves $u^1 =$ const. are the orthogonal trajectories to these geodesics. Figure 4.2 shows the idea.

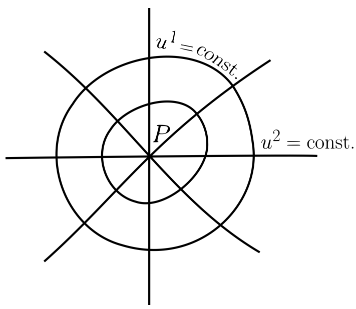


Figure 4.2: The geodesic polar coordinates with centre at P. Why the curves $u^1 =$ const. are called the *geodesic circles* should be clear.

In the proof of the Gauss-Bonnet theorem we will shrink a geodesic circle to the point P, and we are therefore interested in the limiting behaviour of the first fundamental form as we approach P. The coordinates u^1 and u^2 are geodesic parallel coordinates by construction and hence the first fundamental form is given on the form (4.4.6). From this we have $g_{11} = 1$ and $g_{12} = 0$, but the behaviour of g_{22} at P requires further study. Working with (4.4.7) we find

$$dv^1 = \cos u^2 du^1 - u^1 \sin u^2 du^2, \quad dv^2 = \sin u^2 du^1 + u^1 \cos u^2 du^2,$$

and so the first fundamental form becomes

$$\begin{aligned}
&\hat{g}_{11}(dv^1)^2 + 2\hat{g}_{12} dv^1 dv^2 + \hat{g}_{22}(dv^2)^2 = \\
&\hat{g}_{11}\left(\cos^2 u^2 (du^1)^2 - 2\cos^2 u^1 \sin^2 u^2 du^1 du^2 + (u^1)^2 \sin^2 u^2 (du^2)\right) + \\
&2\hat{g}_{12}\left(\cos u^2 \sin u^2 (du^1)^2 + u^1 \cos^2 u^2 du^1 du^2 - \right.\\
&\qquad \left. u^1 \sin^2 u^2 du^1 du^2 - (u^1)^2 \cos u^2 \sin u^2 (du^2)^2\right) + \\
&\hat{g}_{22}\left(\sin^2 u^2 (du^1)^2 + 2\sin^2 u^1 \cos^2 u^2 du^1 du^2 + (u^1)^2 \cos^2 u^2 (du^2)\right).
\end{aligned}$$

Collecting the coefficients in front of $(du^2)^2$ gives

$$g_{22} = (u^1)^2(\hat{g}_{11} \sin^2 u^2 - 2\hat{g}_{12} \cos u^2 \sin u^2 + \hat{g}_{22} \cos^2 u^2),$$

from which if follows that

$$\lim_{u^1 \to 0} g_{22} = 0. \tag{4.4.8}$$

Since it is needed later we also want to look at

$$\lim_{u^1 \to 0} \frac{\partial \sqrt{g_{22}}}{\partial u^1}. \tag{4.4.9}$$

Recall that the geodesic equations in coordinates v^1, v^2 are $\ddot{v}^r + \hat{\Gamma}^r_{ij} \dot{v}^1 \dot{v}^2 = 0$. Since v^i can be written as $c^i s$ close to P, we find $\dot{v}^i = c^i$ and $\ddot{v}^i = 0$ so that $\hat{\Gamma}^r_{ij} c^i c^j = 0$. This holds for all values of c^i, so the Christoffel symbols must be zero at P. From (4.1.2) it follows in turn that all partial derivatives $\partial \hat{g}_{ij}/\partial u^r$ must vanish at P.

Returning to (4.4.9) we find

$$\begin{aligned}
&\lim_{u^1 \to 0} \frac{\partial \sqrt{g_{22}}}{\partial u^1} = \\
&\lim_{u^1 \to 0} \frac{1/2}{\sqrt{g_{22}}} \left(2u^1 + \frac{\partial}{\partial u^1}(\hat{g}_{11} \sin^2 u^2 - 2\hat{g}_{12} \cos u^2 \sin u^2 + \hat{g}_{22} \cos^2 u^2)\right) = \\
&\lim_{u^1 \to 0} \frac{u^1}{\sqrt{g_{22}}} + 0 = \\
&\lim_{u^1 \to 0} \left(\hat{g}_{11} \sin^2 u^2 - 2\hat{g}_{12} \cos u^2 \sin u^2 + \hat{g}_{22} \cos^2 u^2\right)^{-1/2}.
\end{aligned}$$

Now, as u^1 shrinks to P, $\hat{g}_{ij} \to \delta_{ij}$ and the above formula gives

$$\lim_{u^1 \to 0} \frac{\partial \sqrt{g_{22}}}{\partial u^1} = (\sin^2 u^2 + \cos^2 u^2)^{-1/2} = 1. \tag{4.4.10}$$

4.5 The Gauss-Bonnet Theorem

In this section, we prove the Gauss-Bonnet theorem. There are many versions of the theorem. We do not present the most famous version, involving the Euler characteristic, which connects the geometry of a surface to its topology. We are concerned with the local version, for which the proof is slightly involved. With the general theorem in place, we explore a strategy to answer a question posed in the introduction: Can a two-dimensional being on a surface figure out the geometry of its world?

Theorem 10 (Gauss-Bonnet). *Let S be a simply connected portion of a surface M represented by the smooth surface patch $\mathbf{X}(u^1, u^2)$ whose boundary C is a simple closed curve with representation $\mathbf{X}(u^1(s), u^2(s))$ where s is the arc length of C. Let k_g be the geodesic curvature of this curve and K be the Gaussian curvature of S. Then*

$$\int_C k_g ds + \iint_S K dA = 2\pi, \tag{4.5.1}$$

where dA is the element of area of S and the integration along C is carried out such that S stays on the left side.

Proof. The outline of this proof is in three parts. First, we introduce orthogonal coordinates u^1, u^2 on the surface. Then, we pick an arbitrary point and make a small geodesic circle around it. At last, we observe what happens to the integrals in the theorem as the circle reduces to the point.

Let us introduce allowable orthogonal coordinates u^1, u^2. The existence of such coordinates will be covered as a remark to the theorem. Let P be an arbitrary point on S and let C_0 be a sufficiently small geodesic circle with centre at P so that u^1, u^2 are polar geodesic coordinates inside C_0. We also assume that for the boundary curve C we have $u^1 = 1$ so that along this curve u^2 is equal to the arc length s.

Since we have orthogonal coordinates and $c^1 = 1$ on C, equation (4.4.3) gives

$$k_g(C) = \frac{1}{2g_{22}\sqrt{g_{11}}}\frac{\partial g_{22}}{\partial u^1} = \frac{1}{\sqrt{g_{11}g_{22}}}\left(\frac{1}{2\sqrt{g_{22}}}\frac{\partial g_{22}}{\partial u^1}\right) = \frac{1}{\sqrt{g_{11}g_{22}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1}$$

and since, in addition, on this curve $u^2 = 1$, we have $g_{22} = 1$. This gives

$$\int_C k_g ds = \int_C \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1} du^2. \tag{4.5.2}$$

Let S' denote the region bounded by the two curves C_0 and C. Let us study the surface integral over this region. With $dA = \sqrt{g} du^1 du^2$ and the help of (4.4.1) we find

$$\iint_{S'} KdA = -\iint_{S'} \frac{\partial}{\partial u^1}\left(\frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1}\right) du^1du^2$$
$$-\iint_{S'} \frac{\partial}{\partial u^2}\left(\frac{1}{\sqrt{g_{22}}}\frac{\partial\sqrt{g_{11}}}{\partial u^2}\right) du^1du^2.$$

The second integral in this expression is zero since integration with respect to u^2 is along closed curves. For the first integral we find

$$\iint_{S'} KdA = -\int_C \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1}du^2 + \int_{C_0} \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1}du^2.$$

We recognise the first term here as the total geodesic curvature along C as in (4.5.2), and so

$$\int_C k_g ds + \iint_{S'} KdA = \int_{C_0} \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1}du^2.$$

Let us now take the limit as C_0 shrinks to the point P. Then, the integral over the region S' goes to S and

$$\int_C k_g ds + \iint_{S} KdA = \lim_{C_0\to P}\int_{C_0} \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1}du^2.$$

In the region inside C_0, we had by construction geodesic polar coordinates, for which $g_{11} = 1$. In addition (4.4.10) gives

$$\lim_{C_0\to P}\int_{C_0} \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1}du^2 = \lim_{C_0\to P}\int_{C_0} du^2 = 2\pi.$$

□

Remark. The existence of the orthogonal coordinate system used in the theorem is not obvious. One way to show it is to find a conformal mapping from our simply connected portion of a surface to a region where orthogonal coordinates are easy to introduce. A conformal mapping is just another name for angle-perserving,[2] so in particular orthogonal coordinates under a conformal mapping stay orthogonal. The map we want can be found in two steps. First we use Theorem 62.1 in [10] which states that 'Any simply-connected portion of a surface which has a representation of class $r \geq 3$ can be conformally mapped into a plane.' Our smooth surface patch is indeed 'of class $r \geq 3$' so this applies here. Let S^*, C^* and C_0^* be the images of S, C and C_0 under this mapping.

Now that our region is in the plane, we can rely on results for analytic functions in complex analysis. Consulting the classic text *Complex Analysis*

[2]The name conformal was introduced by German astronomer and geographer Schubert as early as 1789 [18].

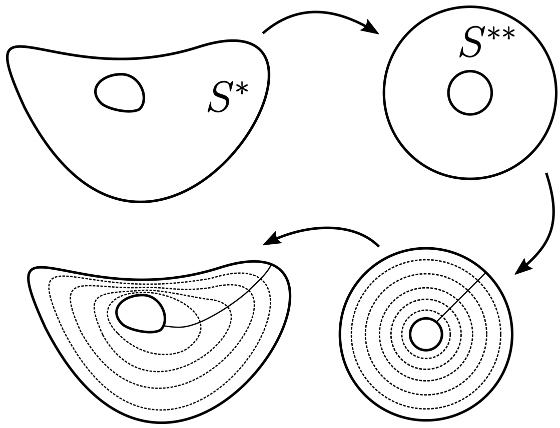


Figure 4.3: The conceptual idea of the existence argument for the orthogonal coordinates. This shows only the mapping from S^* to S^{**} which is based on a theorem from complex analysis. Notice how in the sketch the connected line on the lower part stays orthogonal to the dotted lines.

by Lars V. Ahlfors [19], we find in chapter 6.5 the result we want. His Theorem 10 assures us that there exists a one-to-one conformal mapping from the region S^* to some annulus S^{**}. Figure 4.3 shows the idea. On the annulus, it is easy to introduce orthogonal coordinates by taking concentric circles that cover S^{**} and their orthogonal trajectories. Taking the inverse images of the two mappings results in orthogonal curves on S, if we in addition use the geodesic circles and their orthogonal geodesics inside C_0.

If the curve C is simply connected and continuous but has finitely many corners, we can still use the Gauss-Bonnet theorem after a slight modification. A corner is a discontinuity in the derivative, with P in Figure 4.4 being an example. In a neighbourhood of the point P we replace the curve C by a portion of a geodesic circle of radius r, which we will call C_g. We choose r such that for two points Q and R close to P the tangents of the two curves C and C_g are equal, see Figure 4.4a.

As Q and R approach P, the geodesic circle approaches a perfect circle of radius r. Now, call a length element of the geodesic circle by $d\hat{s}$. We find that

$$\lim_{r\to 0}\int_{C_g} k_g d\hat{s} = \lim_{r\to 0}\int_{C_r} \frac{1}{r} r d\theta = \alpha$$

where we have used that the curvature of a circle is $1/r$ and the length element is $rd\theta$. In this expression, C_r denotes the circular arc from Q to R. The angle α is shown in Figure 4.5 and is the angle between the two tangents at P turning in the direction of the curve C. The interior angle $\beta = \pi - \alpha$ is also shown in the same figure.

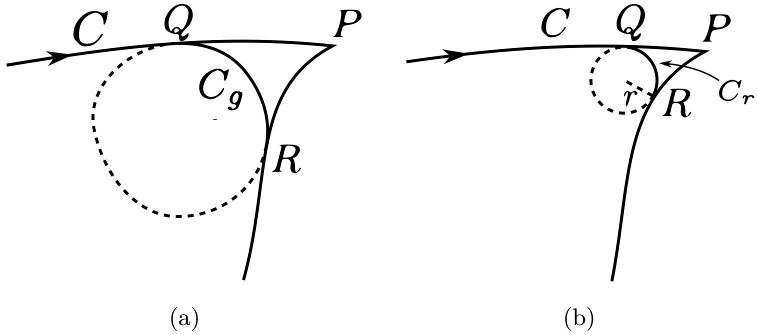


Figure 4.4: (a) and (b) depicts the same corner P of the curve C. At the points Q and R the tangent of the geodesic circle C_g and the original curve C coincide. In (b) the points have moved closer to the corner and the geodesic circle approaches a perfect circle. The segment of the circle between Q and R we call C_r. For a circle, integration of arc length is trivial.

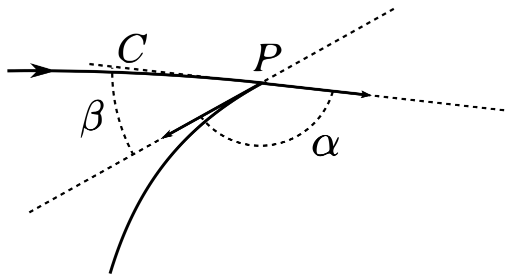


Figure 4.5: The exterior angle α is the directed angle of rotation of the tangent to the curve C at the corner P. The interior angle, $\beta = \pi - \alpha$, is also shown.

Now, assuming that the curve C has n corners $P_1, P_2, \ldots, P_n$ we find a corrected version of the Gauss-Bonnet theorem by replacing total geodesic curvature in (4.5.1) with the geodesic curvature of the line elements *plus* the exterior angles $\alpha_i (i = 1, 2, \ldots n)$. This results in

$$\int_C k_g ds + \sum_{i=1}^{n} \alpha_i + \iint_S K dA = 2\pi,$$

or

$$\int_C k_g ds + \iint_S K dA = \sum_{i=1}^{n} \beta_i - (n-2)\pi, \tag{4.5.3}$$

if we want the interior angles.

From (4.5.3) we find that if the curves between the corners are geodesics, i.e. $k_g = 0$ along C,

$$\iint_S K dA = \sum_{i=1}^{n} \beta_i - (n-2)\pi. \tag{4.5.4}$$

In the case $n = 3$ we have a geodesic triangle, and get

$$\iint_S K dA = \beta_1 + \beta_2 + \beta_3 - \pi. \tag{4.5.5}$$

It is worth stressing that this only holds when the triangle is formed by three geodesics. Equation (4.5.5) is in agreement with the fact from plane geometry saying that the sum of the angles of a triangle always is π, since for a plane any straight line is a geodesic and $K = 0$ everywhere so the integral curvature vanishes. If, however, the curvature is positive at every point, the sum of the angles will be greater than π. An example of this is the sphere. On the sphere of radius R we have the constant Gaussian curvature $1/R^2$ and so the surface area of a geodesic triangle, Δ, is

$$\text{Area}(\Delta) = R^2[(\beta_1 + \beta_2 + \beta_3 - \pi)] = R^2 E.$$

The amount by which the sum of the angles exceeds π is called the spherical excess, E, and was already published by Albert Girard in 1629 [20].

In theory, a two-dimensional object living on any surface could use the excess of a geodesic triangle to measure the Gaussian curvature at a point. Let Δ_k be a geodesic triangle with internal angles $\beta_1^{(k)}, \beta_2^{(k)}$ and $\beta_3^{(k)}$, and containing a point P. The index k is used since we want to create smaller and smaller triangles containing P. See Figure 4.6 for an illustration. Now (4.5.5) gives, dividing by the area of Δ_k,

$$\frac{\iint_{\Delta_k} K dA}{\iint_{\Delta_k} dA} = \frac{\beta_1^{(k)} + \beta_2^{(k)} + \beta_3^{(k)} - \pi}{\iint_{\Delta_k} dA}.$$

Let us now make a sequence $(\Delta_k)_{k \in \mathbb{N}}$ such that for each k the area of Δ_{k+1} is less than the area of Δ_k. We require that all triangles are geodesic triangles and contain the point P. In the limit as k increases to infinity, the geodesic triangle reduces to P, and since the Gaussian curvature varies continuously

$$\lim_{k \to \infty} \frac{\iint_{\Delta_k} K dA}{\iint_{\Delta_k} dA} = K(P)$$

and

$$K(P) = \lim_{k \to \infty} \frac{\beta_1^{(k)} + \beta_2^{(k)} + \beta_3^{(k)} - \pi}{\iint_{\Delta_k} dA}.$$

The effect of this is that, assuming you could measure angles and areas, by forming smaller and smaller geodesic triangles around a point P you could make a sequence of numbers that would in the limit approximate the Gaussian curvature at P. In popular terms: A two-dimensional being on a surface can figure out the geometry of its world.

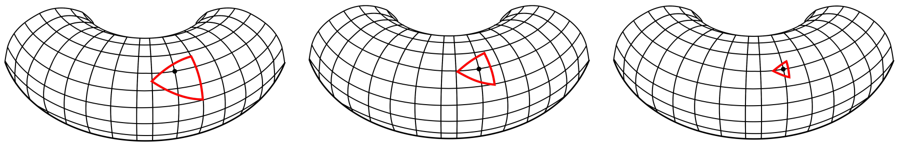


Figure 4.6: From left to right: Examples of smaller and smaller geodesic triangles around a point on a surface. This could be used to measure the Gaussian curvature at the point.

Chapter 5

Example: The Catenoid

In this chapter we explore the catenoid in some detail. A catenoid is a surface obtained from rotating a catenary curve around its directrix. We will compute its fundamental forms and curvatures, and discuss its geodesics. The catenoid may be parameterised by the patch $\mathbf{X} : U \to \mathbb{R}^3$:

$$\mathbf{X}(u, v) = (a \cosh(\frac{v}{a}) \cos(u), a \cosh(\frac{v}{a}) \sin(u), v), \tag{5.0.1}$$

where $a > 0$ is a constant and we, for simplicity, take $0 < u < 2\pi$ and $-1 < v < 1$. The angle u and height v is shown in Figure 5.1 below. We use g_{ij} to denote the coefficients of the first fundamental form in this chapter, but use $(u, v) = (u^1, u^2)$ for the coordinates.

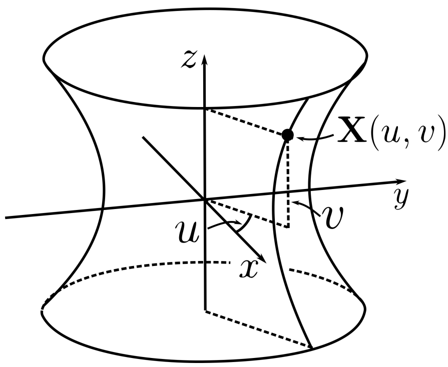


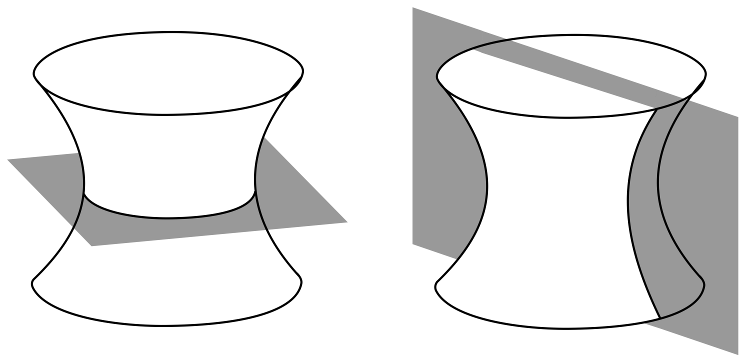
Figure 5.1: The catenoid given by (5.0.1)

5.1 Geodesics on a v-Clairaut Patch

As stated earlier, it is in general difficult to find solutions to the geodesic equations (4.2.4). There are, however, surfaces where finding solutions reduces to computing integrals. The catenoid is such an example. The catenoid is a surface of revolution, for which we have the general form

$$\mathbf{X}(u, v) = (f(v) \cos(u), f(v) \sin(u), g(v)), \tag{5.1.1}$$

for functions $f(v)$ and $g(v)$. Let us assume that $f(v)$ is positive, so that it can be thought of as the radius of the parallel $u \mapsto (f(v)\cos(u), f(v)\sin(u), g(v))$. We also have the meridian where the angle u is fixed. The idea of this is shown in Figure 5.2.



(a) (b)

Figure 5.2: The catenoid is a surface of revolution. In (a) a parallel is shown and (b) depicts a meridian. On such a surface the meridians can be parameterised as geodesics. This can we understand geometrically. Since any meridian would be a plane curve in the plane spanned by the surface normal $\mathbf{N}$ and the axis of rotation, its acceleration only happen in the normal direction which is a characteristic of a geodesic. When parallels (a) are geodesics is discussed below.

Let us compute the coefficients of the first fundamental form of the general surface of revolution (5.1.1). We find

$$\begin{aligned} g_{11} &= \mathbf{X}_1 \cdot \mathbf{X}_1 = f^2(v), \\ g_{12} &= \mathbf{X}_1 \cdot \mathbf{X}_2 = 0, \\ g_{22} &= \mathbf{X}_2 \cdot \mathbf{X}_2 = (f'(v))^2 + (g'(v))^2. \end{aligned}$$

Since $g_{12} = 0$, we have orthogonal coordinates and can use what we found in the section 4.4. When, in addition to $g_{12} = 0$, the relations

$$\frac{\partial g_{11}}{\partial u} = \frac{\partial g_{22}}{\partial u} = 0$$

hold, we have what is called a $v-$Clairaut patch [14]. This simplifies the Christoffel symbols and thus gives us a better chance at solving the geodesic equations. Doing the calculations, one can show that for a $v-$Clairaut patch the Christoffels symbols are

$$\begin{aligned}
\Gamma^1_{11} &= 0, & \Gamma^2_{11} &= \frac{-1}{2g_{22}}\frac{\partial g_{11}}{\partial v}, \\
\Gamma^1_{12} &= \frac{1}{2g_{11}}\frac{\partial g_{11}}{\partial v}, & \Gamma^2_{12} &= 0, \\
\Gamma^1_{22} &= 0, & \Gamma^2_{22} &= \frac{1}{2g_{22}}\frac{\partial g_{22}}{\partial v}.
\end{aligned}$$

These formulas give, by (4.2.4), the geodesic equations

$$\ddot{u} + g^{11}\frac{\partial g_{11}}{\partial v}\dot{u}\dot{v} = 0 \tag{5.1.2}$$

and

$$\ddot{v} - \frac{1}{2}g^{22}\frac{\partial g_{11}}{\partial v}\dot{u}^2 + \frac{1}{2}g^{22}\frac{\partial g_{22}}{\partial v}\dot{v}^2 = 0. \tag{5.1.3}$$

As indicated in the Figure 5.2 we can consider two types of geodesics directly without solving the geodesic equations. For a coordinate curve, we have either a meridian or a parallel. For $u = u_0 = \text{const.}$ we found in (4.4.3) that

$$k_g = \frac{1}{2g_{22}\sqrt{g_{11}}}\frac{\partial g_{22}}{\partial u} \qquad \text{when } u = \text{const. and } g_{12} = 0.$$

Since $\frac{\partial g_{22}}{\partial u} = 0$ for the type of patch we consider, the meridians are thus geodesics. A direct geometric interpretation is given in caption of Figure 5.2. Similarly, for $v = v_0 = \text{const.}$ we found in (4.4.5) that

$$k_g = -\frac{1}{2g_{11}\sqrt{g_{22}}}\frac{\partial g_{11}}{\partial v} \qquad \text{when } v = \text{const. and } g_{12} = 0.$$

This shows that a parallel is a geodesic if and only if $\frac{\partial g_{11}}{\partial v}(u, v_0)$ is zero along the curve $u \mapsto \mathbf{X}(u, v_0)$.

For the remaining curves, the geodesics equations reduce to solving an integral. We adopt Corollary 18.29 in [14], but give a direct proof.

Proposition 6. *Let $\mathbf{X}$ be a v-Clairout patch for the surface M. A curve $\boldsymbol{\alpha} : (a, b) \to M$ of the form*

$$\boldsymbol{\alpha}(v) = \mathbf{X}(u(v), v)$$

is a pregeodesic if and only if there is a constant c such that

$$\frac{du}{dv} = \pm c\sqrt{\frac{g_{22}}{g_{11}(g_{11} - c^2)}}. \tag{5.1.4}$$

Remark. A pregeodesic is just a reparameterisation of a geodesic.

Proof. To prove this, we make the following observation:

$$\begin{aligned}\frac{d}{ds}(g_{11}\dot{u}) =& \frac{dg_{11}}{ds}\dot{u} + g_{11}\ddot{u}\\ =& \frac{\partial g_{11}}{\partial v}\dot{v}\dot{u} + \frac{\partial g_{11}}{\partial u}\dot{u}^2 + g_{11}\ddot{u}.\end{aligned}$$

Since $\partial g_{11}/\partial u = 0$ for a v-Clairout patch, we are left with

$$\frac{d}{ds}(g_{11}\dot{u}) = \frac{\partial g_{11}}{\partial v}\dot{v}\dot{u} + g_{11}\ddot{u} = 0$$

where the last equality comes from the geodesic equations (5.1.2). This shows that

$$g_{11}\dot{u} = c = \text{const.} \tag{5.1.5}$$

Take now a unit-speed curve $\boldsymbol{\alpha}(s) = \mathbf{X}(u(s), v(s))$ on the surface patch , i.e.

$$1 = \|\dot{\boldsymbol{\alpha}}\|^2 = g_{11}\dot{u}^2 + 2g_{12}\dot{u}\dot{v} + g_{11}.$$

Using that $g_{12} = 0$ and (5.1.5) we find

$$\frac{c^2}{g_{11}} + g_{22}\dot{v}^2 = 1$$

or

$$\dot{v} = \pm\sqrt{\frac{1 - c^2/g_{11}}{g_{22}}}.$$

Now, we can combine the equations for $\dot{u}$ and $\dot{v}$ to find

$$\frac{\dot{u}}{\dot{v}} = \frac{du}{dv} = \pm c\sqrt{\frac{g_{22}}{g_{11}(g_{11} - c^2)}}.$$

□

5.2 Fundamental Forms and Curvatures

With all of this in place, let us return to the catenoid. Comparing (5.0.1) and (5.1.1) we find:

$$\begin{aligned}f(v) &= a\cosh(\frac{v}{a}),\\ g(v) &= v.\end{aligned}$$

The coefficients of the first fundamental form are thus

$$g_{11} = f^2(v) = a^2\cosh^2(\frac{v}{a})$$

and

$$g_{22} = (f'(v))^2 + (g'(v))^2 = \sinh^2(\frac{v}{a}) + 1 = \cosh^2(\frac{v}{a}).$$

For the second fundamental form we need the normal vector $\mathbf{N}$ and compute

$$\mathbf{X}_1 \times \mathbf{X}_2 = \left(a \cosh(\frac{v}{a}) \cos(u), a \cosh(\frac{v}{a}) \sin(u), -a \cosh(\frac{v}{a}) \sinh(\frac{v}{a})\right)$$

so that by (2.4.1)

$$\mathbf{N} = \frac{\mathbf{X}_1 \times \mathbf{X}_2}{\|\mathbf{X}_1 \times \mathbf{X}_2\|} = \left(\operatorname{sech}(\frac{v}{a}) \cos(u), \operatorname{sech}(\frac{v}{a}) \sin(u), -\operatorname{sech}(\frac{v}{a}) \sinh(\frac{v}{a})\right).$$

This gives us the chance to compute the coefficients of the second fundamental form by equation (3.2.6):

$$\begin{aligned} b_{11} &= -a, \\ b_{12} &= 0, \\ b_{22} &= \frac{1}{a}. \end{aligned}$$

The Gaussian curvature is, by Lemma 2,

$$K = \frac{b}{g} = \frac{-1}{a^2 \cosh^4(\frac{v}{a})}.$$

For the principal curvatures, we showed earlier that they were the roots of the determinant of the matrix (3.3.4):

$$\begin{bmatrix} b_{11} - k_n g_{11} & b_{12} - k_n g_{12} \\ b_{12} - k_n g_{12} & b_{22} - k_n g_{22} \end{bmatrix} = \begin{bmatrix} -a - k_n a^2 \cosh^2(\frac{v}{a}) & 0 \\ 0 & 1/a - k_n \cosh^2(\frac{v}{a}) \end{bmatrix}.$$

From this, we compute

$$k_1 = \frac{1}{a \cosh^2(\frac{v}{a})} \quad \text{and} \quad k_2 = \frac{-1}{a \cosh^2(\frac{v}{a})},$$

and hence find a vanishing mean curvature

$$H = \frac{1}{2}(k_1 + k_2) = 0.$$

The vanishing mean curvature is sufficient for a surface to be a minimal surface [14].

5.3 The Geodesic Curves

Let us, at last, consider the geodesic curves on the catenoid. First, we have that the meridians for a fixed value of the angle u are geodesics. Second, for the parallels to be geodesics we needed $\frac{\partial g_{11}}{\partial v}$ to vanish for all values of u. We find

$$\frac{\partial g_{11}}{\partial v} = a \sinh(\frac{2v}{a}) = 0 \text{ for } v = 0,$$

and so only the parallel at the thinnest point of the catenoid is a geodesic. This is in agreement with our intuition since at any other height the parallel curve must curve within the surface to stay on path. Or, in other words, $v = 0$ is the only height for which the acceleration of the parallel curve is exclusively in the direction of the surface normal $\mathbf{N}$.

For the other geodesics we use (5.1.4) to find

$$\frac{du}{dv} = \pm c \sqrt{\frac{\cosh^2(v/a)}{a^2 \cosh^2(v/a)(\cosh^2(v/a) - c^2)}} = \frac{\pm c/a}{\sqrt{\cosh^2(v/a) - c^2}}. \quad (5.3.1)$$

Integrating this to find the function $u = u(v)$ does not come without difficulties, and looking for the correct reduction to an elliptic integral we can solve (e.g. in the *Handbook of Elliptic Integrals for Engineers and Scientists* [21]) is outside the scope of this thesis. However, since we are curious as to what the geodesics may look like we had a go at the geodesics going through the point $(u, v) = (0, 0)$. Ignoring all possible problems with the integral an attempt at solving it numerically for $a = 1$ and for a few values of $c < 1$ is shown in Figure 5.3. There, we also include the meridian and parallel.

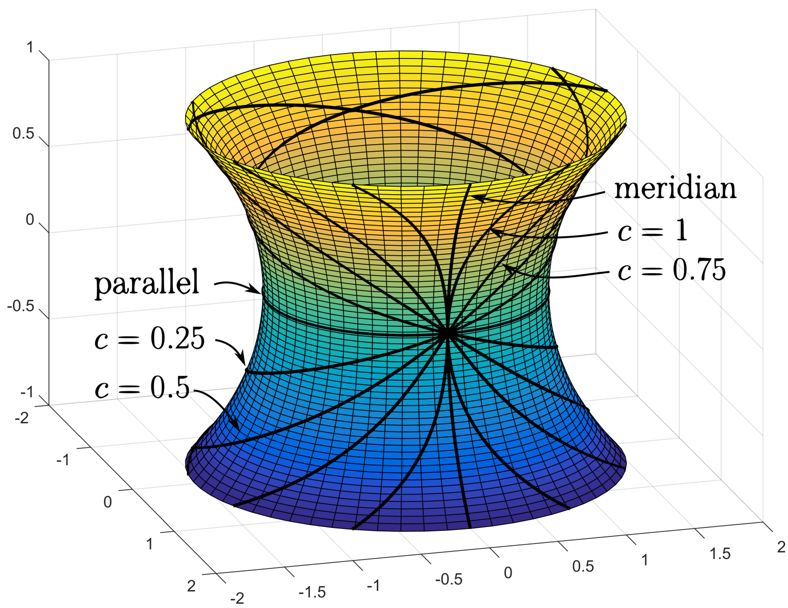


Figure 5.3: Some geodesics on the catenoid (5.0.1) going through the point $(u, v) = (0, 0)$, for $a = 1$. The numerical integration was carried out in Matlab with using the ODE solver ode45 (see [22]) for different values of c.

Chapter 6

Riemannian Geometry

We end this thesis returning to some historical aspects of the theory of surfaces and show that a lot of the classic results we have demonstrated transfer into the abstract setting of Riemannian geometry. Let us begin on a historical note.

Bernhard Riemann held in 1854 a famous lecture ‘On the hypotheses which lie at the bases of geometry’. It was a mathematical lecture on geometry, with no illustrations and with barely any formulas, which today is recognised as one of the most important contributions to mathemathics. On why it had such a wide effect, Jürgen Jost writes in a book dedicated to the lecture [23]:

> ‘This is because its position is at the intersection of mathematics, physics and philosophy, and it not only founds and establishes a central mathematical discipline, but also paves way for the physics of the twentieth century and at the same time represent a timeless refutation of certain philosophical concepts of space.’

Riemann’s teacher, Gauss, had at the time discovered interesting connections between intrinsic and extrinsic properties of surfaces. In particular, the remarkable result that curvature could be intrinsic. Inspired by this, Riemann took in his work an abstract and intrinsic view of geometry, a geometry no longer confined to objects in Euclidean space. This approach was fruitful and, in addition to vast developments in mathematics, the theory was later essential to Einstein’s celebrated work connecting gravity to the geometry of space and time. We now try to connect this abstract approach to the results shown in the previous chapters.

In Riemannian geometry, we study manifolds[1] of arbitrary dimension, n, given by local coordinates $\mathbf{u} = (u^1, u^2, \dots, u^n)$ where $\mathbf{u} \in U \subset \mathbb{R}^n$. As pointed out in the remark following Definition 13, there are different ways of viewing tangent vectors. Since we want to remove the connection to Euclidean space, we no longer define tangent vectors as vectors in Euclidean space. A better way is to view tangent vectors as differential operators.

With this abstraction, we can define an almost arbitrary inner product on the tangent space, i.e. between tangent vectors. Riemannian geometry considers inner products that vary smoothly from point to point. The inner

[1]A concept we do not define, but note that our definitions of a surface in section 2.2 by patches, atlases, transitions maps and homeomorphisms are key components.

product allows us, as before, to discuss angles, volumes, lengths of curves, and more. With reference to a patch on the manifold with local coordinates $\mathbf{u}$, two tangent vectors can be considered as $\mathbf{v} = v^i \boldsymbol{\partial}/\boldsymbol{\partial}\mathbf{u^i}$ and $\mathbf{w} = w^i \boldsymbol{\partial}/\boldsymbol{\partial}\mathbf{u^i}$ with inner product

$$\langle \mathbf{v}, \mathbf{w} \rangle = g_{ij}(\mathbf{u}) v^i w^j.$$

A lot of the results concerning curves on surfaces extends to this higher dimensional abstract setting. The speed of the curve in a manifold M, $\boldsymbol{\alpha} : (a, b) \to M$, at $t \in (a, b)$ is $\|\boldsymbol{\alpha}'(t)\| = \langle \boldsymbol{\alpha}'(t), \boldsymbol{\alpha}'(t) \rangle^{1/2}$ and, with s being arc length,

$$ds^2 = g_{ij} du^i du^j.$$

The geodesic equations (4.2.4)

$$\ddot{u}^r + \Gamma^r_{ij} \dot{u}^i \dot{u}^j = 0, \quad r = 1, 2, \ldots, n$$

still describe locally length-minimising 'straight' curves. This becomes apparent if we go back to the proof provided in section 4.3 and realise that it is intrinsic, and easily extends to higher dimensions. The same can be said about Theorem 9 and 10, and also the results of section 4.4. The Christoffel symbols, originally defined by the Gauss formulas (3.2.2), can now take (4.1.2) as definition.

If we restrict ourself to two-dimensional manifolds of the above structure, called Riemannian surfaces, we can define the Gaussian curvature directly, e.g. by (3.3.5). However, not all concepts can be discussed in this setting. Talking about normal directions, principal curvatures and mean curvature does not make sense. Here we still need an embedding in Euclidean space.

Riemann did more than what we have outlined here, and later Beltrami, Bianchi, Levi-Civita, Ricci and others further developed the theory into modern Riemannian geometry where connections and parallel transport are central [24]. As an introduction to the modern theory, the book *Riemannian Geometry and Geometric Analysis* by Jürgen Jost [25] looks like a natural next step.

References

[1] Scriba CJ, Schreiber P. 5000 Years of Geometry. Basel: Birkhäuser; 2015.

[2] Boyer CB, Merzbach UC. A History of Mathematics. 2nd ed. New York: Wiley; 1989.

[3] Euler L. Recherches sur la courbure des surfaces. Memoires de l'academie des sciences de Berlin. 1767;16:119–143.

[4] Gauss CF. Disquisitiones generales circa superficies curvas. Göttingen: Dieterich; 1828.

[5] Joyce DE. Euclids Elements. Book 1 Postulate 5 [Internet]; [Updated: 2003. Accessed: 31 May 2016]. Available from: http://aleph0.clarku.edu/~djoyce/elements/bookI/post5.html.

[6] Inkscape. ver. 0.91 [Computer Software]; 2015. Available from: https://inkscape.org/.

[7] The MathWorks Inc . MATLAB ver. 8.6.0 (R2015b) [Computer Software]. Natick, Massachusetts; 2015.

[8] Pressley A. Elementary Differential Geometry. 1st ed. London: Springer-Verlag; 2001.

[9] Faber RL. Differential Geometry and Relativity Theory. 1st ed. New York: Marcel Dekker, Inc.; 1983.

[10] Kreyszig E. Differential Geometry. Dover Books on Mathematics. New York: Dover Publications; 1991.

[11] Jordan D, Smith P. Nonlinear Ordinary Differential Equations. 4th ed. Oxford University Press; 2007.

[12] Willmore TJ. An Introduction to Differential Geometry. Oxford: Clarendon Press; 1959.

[13] Isham CJ. Modern Differential Geometry for Physicists. 2nd ed. World Scientific Lecture Notes in Physics. Singapore: World Scientific; 1999.

[14] Gray A, Abbena E, Salamon S. Modern Differential Geometry of Curves and Surfaces with Mathematica. 3rd ed. Studies in Advanced Mathematics. Boca Raton: Chapman & Hall/CRC; 2006.

[15] Christoffel EB. Über die Transformation der homogenen Differentialausdrücke zweiten Grades. Journal für die reine und angewandte Mathematik. 1869;70:46–70. Available from: http://eudml.org/doc/148073.

[16] Jost J, Li-Jost X. Calculus of Variations. Cambridge studies in advanced mathematics 64. Cambridge University Press; 1998.

[17] Kelley WG, Peterson AC. Theory of Differential Equations. 2nd ed. Universitext. New York: Springer-Verlag; 2010.

[18] Schubert TV. De proiectione Sphaeroidis ellipticae geographica. Nova Acta Academiae Scientiarum Imperialis petropolitana. 1789;5:130–146.

[19] Ahlfors LV. Complex Analysis. 3rd ed. International Series in Pure and Applied Mathematics. New York: McGraw-Hill; 1979.

[20] Girard A. Invention nouvelle en l'algèbre. Blauew; 1629. Available from: http://gallica.bnf.fr/ark:/12148/bpt6k5822034w.

[21] Byrd PF, Friedman MD. Handbook of Elliptic Integrals for Engineers and Scientists. Berlin: Springer-Verlag Heidelberg; 1971.

[22] The MathWorks Inc . Matlab R2015b Documentation: ode45 [Internet]; 2015. [Accessed: 1 June 2016]. Available from: http://se.mathworks.com/help/matlab/ref/ode45.html.

[23] Riemann B. On the Hypotheses Which Lie at the Bases of Geometry. Jost J, editor. Classic Texts in the Sciences. Basel: Birkhäuser; 2016.

[24] Heckman G. Classical Differential Geometry [Lecture Notes] Radboud University Nijmegen; 2016. Available from: http://www.math.ru.nl/~heckman/CDG.pdf.

[25] Jost J. Riemannian Geometry and Geometric Analysis. 6th ed. Universitext. Berlin: Springer-Verlag Heidelberg; 2011.