Chapter 20
Basics of the Differential Geometry of Surfaces

20.1 Introduction

The purpose of this chapter is to introduce the reader to some elementary concepts of the differential geometry of surfaces. Our goal is rather modest: We simply want to introduce the concepts needed to understand the notion of Gaussian curvature, mean curvature, principal curvatures, and geodesic lines. Almost all of the material presented in this chapter is based on lectures given by Eugenio Calabi in an upper undergraduate differential geometry course offered in the fall of 1994. Most of the topics covered in this course have been included, except a presentation of the global Gauss–Bonnet–Hopf theorem, some material on special coordinate systems, and Hilbert's theorem on surfaces of constant negative curvature.

What is a surface? A precise answer cannot really be given without introducing the concept of a manifold. An informal answer is to say that a surface is a set of points in $\mathbb{R}^3$ such that for every point p on the surface there is a small (perhaps very small) neighborhood U of p that is continuously deformable into a little flat open disk. Thus, a surface should really have some topology. Also, locally, unless the point p is "singular," the surface looks like a plane.

Properties of surfaces can be classified into *local properties* and *global properties*. In the older literature, the study of local properties was called *geometry in the small*, and the study of global properties was called *geometry in the large*. Local properties are the properties that hold in a small neighborhood of a point on a surface. Curvature is a local property. Local properties can be studied more conveniently by assuming that the surface is parametrized locally. Thus, it is important and useful to study parametrized patches. In order to study the global properties of a surface, such as the number of its holes or boundaries, global topological tools are needed. For example, closed surfaces cannot really be studied rigorously using a single parametrized patch, as in the study of local properties. It is necessary to cover a closed surface with various patches, and these patches need to overlap in some clean fashion, which leads to the notion of a manifold.

Another more subtle distinction should be made between *intrinsic* and *extrinsic* properties of a surface. Roughly speaking, intrinsic properties are properties of a surface that do not depend on the way the surface is immersed in the ambient space, whereas extrinsic properties depend on properties of the ambient space. For example, we will see that the Gaussian curvature is an intrinsic concept, whereas the normal to a surface at a point is an extrinsic concept. The distinction between these two notions is clearer in the framework of Riemannian manifolds, since manifolds provide a way of defining an abstract space not immersed in some a priori given ambient space, but readers should have some awareness of the difference between intrinsic and extrinsic properties.

In this chapter we focus exclusively on the study of local properties, both intrinsic and extrinsic, and manifolds are completely left out. Readers eager to learn more differential geometry and about manifolds are refereed to do Carmo [12], Berger and Gostiaux [4], Lafontaine [29], and Gray [23]. A more complete list of references can be found in Section 20.11.

By studying the properties of the curvature of curves on a surface, we will be led to the first and second fundamental forms of a surface. The study of the normal and tangential components of the curvature will lead to the normal curvature and to the geodesic curvature. We will study the normal curvature, and this will lead us to principal curvatures, principal directions, the Gaussian curvature, and the mean curvature. In turn, the desire to express the geodesic curvature in terms of the first fundamental form alone will lead to the Christoffel symbols. The study of the variation of the normal at a point will lead to the Gauss map and its derivative, and to the Weingarten equations. We will also quote Bonnet's theorem about the existence of a surface patch with prescribed first and second fundamental forms. This will require a discussion of the *Theorema Egregium* and of the Codazzi–Mainardi compatibility equations. We will take a quick look at curvature lines, asymptotic lines, and geodesics, and conclude by quoting a special case of the Gauss–Bonnet theorem.

Since this chapter is just a brief introduction to the local theory of the differential geometry of surfaces, the following additional references are suggested. For an intuitive introduction to differential geometry there is no better source that the beautiful presentation given in Chapter IV of Hilbert and Cohn-Vossen [25]. The style is informal, and there are occasional mistakes, but there are amazingly powerful geometric insights. The reader will have a taste of the state of differential geometry in the 1920s. For a taste of the differential geometry of surfaces in the 1980s, we highly recommend Chapter 10 and Chapter 11 in Berger and Gostiaux [4]. These remarkable chapters are written as a guide, basically without proofs, and assume a certain familiarity with differential geometry, but we believe that most readers could easily read them after completing this chapter. For a comprehensive and yet fairly elementary treatment of the differential geometry of curves and surfaces we highly recommend do Carmo [12] and Kreyszig [28]. Another nice and modern presentation of differential geometry including many examples in *Mathematica* can be found in Gray [23]. The older texts by Stoker [42] and Hopf [26] are also recommended. For the (very) perseverant reader interested in the state of surface theory around the 1900s, nothing tops Darboux's four–volume treatise [9, 10, 7, 8]. Actually, Darboux is a real gold mine for all sorts of fascinating (often long forgotten) results. For a very interesting article on the history of differential geometry see Paulette Libermann's article in Dieudonné [11], Chapter IX. More references can be found in Section 20.11. Some interesting applications of the differential geometry of surfaces to geometric design can be found in the Ph.D. theses of Henry Moreton [38] and William Welch [44]; see Section 20.13 for a glimpse of these applications.

20.2 Parametrized Surfaces

In this chapter we consider exclusively surfaces immersed in the affine space $\mathbb{A}^3$. In order to be able to define the normal to a surface at a point, and the notion of curvature, we assume that some inner product is defined on $\mathbb{R}^3$. Unless specified otherwise, we assume that this inner product is the standard one, i.e.,

$$(x_1,x_2,x_3)\cdot(y_1,y_2,y_3) = x_1y_1 + x_2y_2 + x_3y_3.$$

The Euclidean space obtained from $\mathbb{A}^3$ by defining the above inner product on $\mathbb{R}^3$ is denoted by $\mathbb{E}^3$ (and similarly, $\mathbb{E}^2$ is associated with $\mathbb{A}^2$).

Let Ω be some open subset of the plane $\mathbb{R}^2$. Recall that a map $X\colon \Omega \to \mathbb{E}^3$ is C^p*-continuous* if all the partial derivatives

$$\frac{\partial^{i+j}X}{\partial u^i \partial v^j}(u,v)$$

exist and are continuous for all i, j such that $0 \leq i+j \leq p$, and all $(u,v) \in \mathbb{R}^2$. A surface is a map $X\colon \Omega \to \mathbb{E}^3$, as above, where X is at least C^3-continuous. It turns out that in order to study surfaces, in particular the important notion of curvature, it is very useful to study the properties of curves on surfaces. Thus, we will begin by studying curves on surfaces. The curves arising as plane sections of a surface by planes containing the normal line at some point of the surface will play an important role. Indeed, we will study the variation of the "normal curvature" of such curves. We will see that in general, the normal curvature reaches a maximum value κ_1 and a minimum value κ_2. This will lead us to the notion of Gaussian curvature (it is the product $K = \kappa_1\kappa_2$).

Actually, we will need to impose an extra condition on a surface X so that the tangent plane (and the normal) at any point is defined. Again, this leads us to consider curves on X.

A curve C on X is defined as a map $C\colon t \mapsto X(u(t),v(t))$, where u and v are continuous functions on some open interval I contained in Ω. We also assume that the plane curve $t \mapsto (u(t),v(t))$ is regular, that is, that

$$\left(\frac{du}{dt}(t), \frac{dv}{dt}(t)\right) \neq (0,0) \text{ for all } t \in I.$$

For example, the curves $v \mapsto X(u_0, v)$ for some constant u_0 are called *u-curves*, and the curves $u \mapsto X(u, v_0)$ for some constant v_0 are called *v-curves*. Such curves are also called the *coordinate curves*.

We would like the curve $t \mapsto X(u(t), v(t))$ to be a regular curve for all regular curves $t \mapsto (u(t), v(t))$, i.e., to have a well-defined tangent vector for all $t \in I$. The tangent vector $dC(t)/dt$ to C at t can be computed using the chain rule:

$$\frac{dC}{dt}(t) = \frac{\partial X}{\partial u}(u(t), v(t))\frac{du}{dt}(t) + \frac{\partial X}{\partial v}(u(t), v(t))\frac{dv}{dt}(t).$$

Note that

$$\frac{dC}{dt}(t), \quad \frac{\partial X}{\partial u}(u(t), v(t)) \quad \text{and} \quad \frac{\partial X}{\partial v}(u(t), v(t))$$

are vectors, but for simplicity of notation, we omit the vector symbol in these expressions.[1]

It is customary to use the following abbreviations: The partial derivatives

$$\frac{\partial X}{\partial u}(u(t), v(t)) \quad \text{and} \quad \frac{\partial X}{\partial v}(u(t), v(t))$$

are denoted by $X_u(t)$ and $X_v(t)$, or even by X_u and X_v, and the derivatives

$$\frac{dC}{dt}(t), \quad \frac{du}{dt}(t) \quad \text{and} \quad \frac{dv}{dt}(t)$$

are denoted by $\dot{C}(t)$, $\dot{u}(t)$, and $\dot{v}(t)$, or even by $\dot{C}$, $\dot{u}$, and $\dot{v}$. When the curve C is parametrized by arc length s, we denote

$$\frac{dC}{ds}(s), \quad \frac{du}{ds}(s), \quad \text{and} \quad \frac{dv}{ds}(s)$$

by $C'(s)$, $u'(s)$, and $v'(s)$, or even by C', u', and v'. Thus, we reserve the prime notation to the case where the parametrization of C is by arc length.

Note that it is the curve $C\colon t \mapsto X(u(t), v(t))$ that is parametrized by arc length, not the curve $t \mapsto (u(t), v(t))$.

Using this notation $\dot{C}(t)$ is expressed as follows:

$$\dot{C}(t) = X_u(t)\dot{u}(t) + X_v(t)\dot{v}(t),$$

or simply as

$$\dot{C} = X_u\dot{u} + X_v\dot{v}.$$

[1] Also, traditionally, the result of multiplying a vector u by a scalar λ is denoted by λu, with the scalar on the left. In the expressions above involving partial derivatives, the scalar is written on the right of the vector rather on the left. Although possibly confusing, this appears to be standard practice.

Now, if we want $\dot{C} \neq 0$ for all regular curves $t \mapsto (u(t), v(t))$, we must require that X_u and X_v be linearly independent. Equivalently, we must require that the cross product $X_u \times X_v$ be nonnull.

Definition 20.1. A *surface patch* X, for short a *surface* X, is a map $X\colon \Omega \to \mathbb{E}^3$ where Ω is some open subset of the plane $\mathbb{R}^2$ and where X is at least C^3-continuous. We say that the surface X is *regular at* $(u,v) \in \Omega$ if $X_u \times X_v \neq 0$, and we also say that $p = X(u,v)$ is *a regular point of* X. If $X_u \times X_v = 0$, we say that $p = X(u,v)$ is a *singular point of* X. The surface X is *regular on* Ω if $X_u \times X_v \neq 0$, for all $(u,v) \in \Omega$. The subset $X(\Omega)$ of $\mathbb{E}^3$ is called the *trace* of the surface X.

Remark: It often often desirable to define a (regular) surface patch $X\colon \Omega \to \mathbb{E}^3$ where Ω is a *closed* subset of $\mathbb{R}^2$. If Ω is a closed set, we assume that there is some open subset U containing Ω and such that X can be extended to a (regular) surface over U (i.e., that X is at least C^3-continuous).

Given a regular point $p = X(u,v)$, since the tangent vectors to all the curves passing through a given point are of the form $X_u\dot{u} + X_v\dot{v}$, it is obvious that they form a vector space of dimension 2 isomorphic to $\mathbb{R}^2$ called the *tangent space at p*, and denoted by $T_p(X)$. Note that (X_u, X_v) is a basis of this vector space $T_p(X)$. The set of tangent lines passing through p and having some tangent vector in $T_p(X)$ as direction is an affine plane called the *affine tangent plane at p*. Geometrically, this is an object different from $T_p(X)$, and it should be denoted differently (perhaps as $AT_p(X)$?).[2] Nevertheless, we will use the notation $T_p(X)$ like everybody else, but by calling it *tangent plane* instead of *tangent space*, we hope that the potential confusion will be eliminated.

The unit vector

$$\mathbf{N}_p = \frac{X_u \times X_v}{\|X_u \times X_v\|}$$

is called the *unit normal vector at p*, and the line through p of direction $\mathbf{N}_p$ is the *normal line to X at p*. This time, we can use the notation N_p for the line, to distinguish it from the vector $\mathbf{N}_p$.

Example 20.1. Let $\Omega =]-1,1[\times]-1,1[$, and let X be the surface patch defined by

$$x = \frac{2au}{u^2+v^2+1}, \quad y = \frac{2bv}{u^2+v^2+1}, \quad z = \frac{c(1-u^2-v^2)}{u^2+v^2+1},$$

where $a,b,c > 0$. The surface X is a portion of an ellipsoid. Let

$$t \mapsto (t, t^2)$$

be the piece of parabola corresponding to $t \in]-1,1[$. Then we obtain the curve $C(t) = X(t,t^2)$ on the surface X. It is easily verified that the unit normal to the

[2] It would probably be better to denote the tangent space by $\overrightarrow{T}_p(X)$ and the tangent plane by $T_p(X)$, but nobody else does!

surface is

$$\mathbf{N}_{(u,v)} = (2bcu/Z, 2acv/Z, ab(1-u^2-v^2)/Z),$$

where

$$Z^2 = 4b^2c^2u^2 + 4a^2c^2v^2 + a^2b^2(1-u^2-v^2)^2.$$

The portion of ellipsoid X, the curve C on X, some unit normals, and some tangent vectors (for $u = \frac{1}{3}, v = \frac{1}{9}$), are shown in Figure 20.1, for $a = 5$, $b = 4$, $c = 3$.

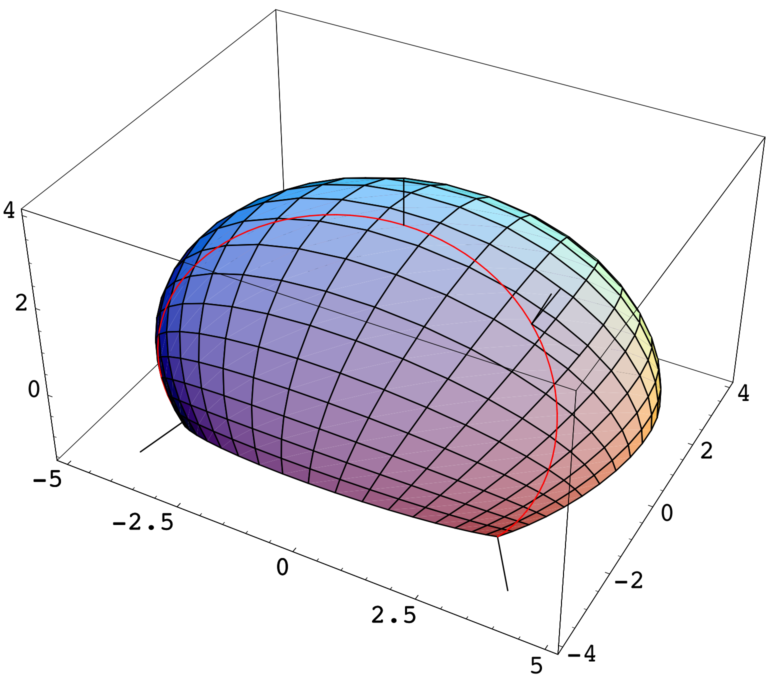

Fig. 20.1 A curve C on a surface X.

The fact that we are not requiring the map X defining a surface $X\colon \Omega \to \mathbb{E}^3$ to be injective may cause problems. Indeed, if X is not injective, it may happen that $p = X(u_0, v_0) = X(u_1, v_1)$ for some (u_0, v_0) and (u_1, v_1) such that $(u_0, v_0) \neq (u_1, v_1)$. In this case, the tangent plane $T_p(X)$ at p is not well-defined. Indeed, we really have two pairs of partial derivatives $(X_u(u_0, v_0), X_v(u_0, v_0))$ and $(X_u(u_1, v_1), X_v(u_1, v_1))$, and the planes spanned by these pairs could be distinct. In this case there are really two tangent planes $T_{(u_0,v_0)}(X)$ and $T_{(u_1,v_1)}(X)$ at the point p where X has a self-intersection. Similarly, the normal $\mathbf{N}_p$ is not well-defined, and we really have two normals $\mathbf{N}_{(u_0,v_0)}$ and $\mathbf{N}_{(u_1,v_1)}$ at p.

We could avoid the problem entirely by assuming that X is injective. This will rule out many surfaces that come up in practice. If necessary, we use the notation

$T_{(u,v)}(X)$ or $\mathbf{N}_{(u,v)}$, which removes possible ambiguities. However, it is a more cumbersome notation, and we will continue to write $T_p(X)$ and $\mathbf{N}_p$, being aware that this may be an ambiguous notation, and that some additional information is needed.

The tangent space may also be undefined when p is not a regular point.

Example 20.2. Considering the surface $X = (x(u,v), y(u,v), z(u,v))$ defined such that

$$\begin{aligned} x &= u(u^2+v^2),\\ y &= v(u^2+v^2),\\ z &= u^2v - v^3/3, \end{aligned}$$

note that all the partial derivatives at the origin $(0,0)$ are zero. Thus, the origin is a singular point of the surface X. Indeed, one can check that the tangent lines at the origin do not lie in a plane.

It is interesting to see how the unit normal vector $\mathbf{N}_p$ changes under a change of parameters. Assume that $u = u(r,s)$ and $v = v(r,s)$, where $(r,s) \mapsto (u,v)$ is a diffeomorphism. By the chain rule,

$$\begin{aligned} X_r \times X_s &= \left(X_u \frac{\partial u}{\partial r} + X_v \frac{\partial v}{\partial r}\right) \times \left(X_u \frac{\partial u}{\partial s} + X_v \frac{\partial v}{\partial s}\right)\\ &= \left(\frac{\partial u}{\partial r}\frac{\partial v}{\partial s} - \frac{\partial u}{\partial s}\frac{\partial v}{\partial r}\right) X_u \times X_v\\ &= \begin{vmatrix} \dfrac{\partial u}{\partial r} & \dfrac{\partial u}{\partial s}\\ \dfrac{\partial v}{\partial r} & \dfrac{\partial v}{\partial s} \end{vmatrix} X_u \times X_v\\ &= \frac{\partial(u,v)}{\partial(r,s)} X_u \times X_v, \end{aligned}$$

denoting the Jacobian determinant of the map $(r,s) \mapsto (u,v)$ by $\partial(u,v)/\partial(r,s)$. Then, the relationship between the unit vectors $\mathbf{N}_{(u,v)}$ and $\mathbf{N}_{(r,s)}$ is

$$\mathbf{N}_{(r,s)} = \mathbf{N}_{(u,v)} \operatorname{sign} \frac{\partial(u,v)}{\partial(r,s)}.$$

We will therefore restrict our attention to changes of variables such that the Jacobian determinant $\partial(u,v)/\partial(r,s)$ is positive.

One should also note that the condition $X_u \times X_v \neq 0$ is equivalent to the fact that the Jacobian matrix of the derivative of the map $X\colon \Omega \to \mathbb{E}^3$ has rank 2, i.e., that the derivative $\mathrm{D}X(u,v)$ of X at (u,v) is injective. Indeed, the Jacobian matrix of the derivative of the map

$$(u,v) \mapsto X(u,v) = (x(u,v), y(u,v), z(u,v))$$

is

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{pmatrix},$$

and $X_u \times X_v \neq 0$ is equivalent to saying that one of the minors of order 2 is invertible. Thus, a regular surface is an *immersion* of an open set of $\mathbb{R}^2$ into $\mathbb{E}^3$.

To a great extent, the properties of a surface can be studied by studying the properties of curves on this surface. One of the most important properties of a surface is its curvature. A gentle way to introduce the curvature of a surface is to study the curvature of a curve on a surface. For this, we will need to compute the norm of the tangent vector to a curve on a surface. This will lead us to the first fundamental form.

20.3 The First Fundamental Form (Riemannian Metric)

Given a curve C on a surface X, we first compute the element of arc length of the curve C. For this, we need to compute the square norm of the tangent vector $\dot{C}(t)$. The square norm of the tangent vector $\dot{C}(t)$ to the curve C at p is

$$\|\dot{C}\|^2 = (X_u\dot{u} + X_v\dot{v}) \cdot (X_u\dot{u} + X_v\dot{v}),$$

where $\cdot$ is the inner product in $\mathbb{E}^3$, and thus,

$$\|\dot{C}\|^2 = (X_u \cdot X_u)\,\dot{u}^2 + 2(X_u \cdot X_v)\,\dot{u}\dot{v} + (X_v \cdot X_v)\,\dot{v}^2.$$

Following common usage, we let

$$E = X_u \cdot X_u, \quad F = X_u \cdot X_v, \quad G = X_v \cdot X_v,$$

and

$$\|\dot{C}\|^2 = E\,\dot{u}^2 + 2F\,\dot{u}\dot{v} + G\,\dot{v}^2.$$

Euler had already obtained this formula in 1760. Thus, the map $(x,y) \mapsto Ex^2 + 2Fxy + Gy^2$ is a quadratic form on $\mathbb{R}^2$, and since it is equal to $\|\dot{C}\|^2$, using the plane curves $t \mapsto (u(t), v(t)) = (xt, yt)$ for any $x, y \in \mathbb{R}$, since $\dot{u} = x$ and $\dot{v} = y$, we show easily that it is positive definite (assuming that $X_u \times X_v \neq 0$). This quadratic form plays a major role in the theory of surfaces, and deserves an official definition.

Definition 20.2. Given a regular surface X, for any point $p = X(u,v)$ on X, letting

$$E = X_u \cdot X_u, \quad F = X_u \cdot X_v, \quad G = X_v \cdot X_v,$$

the positive definite quadratic form $(x,y) \mapsto Ex^2 + 2Fxy + Gy^2$ is called the *first fundamental form of X at p*. It is often denoted by I_p, and in matrix form, we have

$$I_p(x,y) = (x,y) \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Since the map $(x,y) \mapsto Ex^2 + 2Fxy + Gy^2$ is a positive definite quadratic form, we must have $E \neq 0$ and $G \neq 0$. Then, we can write

$$Ex^2 + 2Fxy + Gy^2 = E\left(x + \frac{F}{E}y\right)^2 + \frac{EG - F^2}{E}y^2.$$

Since this quantity must be positive, we must have $E > 0$, $G > 0$, and also $EG - F^2 > 0$.

The symmetric bilinear form φ_I associated with I is an inner product on the tangent space at p, such that

$$\varphi_I((x_1,y_1),(x_2,y_2)) = (x_1,y_1) \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}.$$

This inner product is also denoted by $\langle (x_1,y_1),(x_2,y_2) \rangle_p$. The inner product φ_I can be used to determine the angle of two curves passing through p, i.e., the angle θ of the tangent vectors to these two curves at p. We have

$$\cos\theta = \frac{\langle (\dot{u}_1,\dot{v}_1),(\dot{u}_2,\dot{v}_2) \rangle}{\sqrt{I(\dot{u}_1,\dot{v}_1)}\sqrt{I(\dot{u}_2,\dot{v}_2)}}.$$

For example, the angle between the u-curve and the v-curve passing through p (where u or v is constant) is given by

$$\cos\theta = \frac{F}{\sqrt{EG}}.$$

Thus, the u-curves and the v-curves are orthogonal iff $F(u,v) = 0$ on Ω.

Remarks:

(1) Since

$$\left(\frac{ds}{dt}\right)^2 = \|\dot{C}\|^2 = E\,\dot{u}^2 + 2F\,\dot{u}\dot{v} + G\,\dot{v}^2$$

represents the square of the "element of arc length" of the curve C on X, and since $du = \dot{u}dt$ and $dv = \dot{v}dt$, one often writes the first fundamental form as

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2.$$

Thus, the length $l(pq)$ of an arc of curve on the surface joining the points $p = X(u(t_0), v(t_0))$ and $q = X(u(t_1), v(t_1))$ is

$$l(p,q) = \int_{t_0}^{t_1} \sqrt{E\,\dot{u}^2 + 2F\,\dot{u}\dot{v} + G\,\dot{v}^2}\,dt.$$

One also refers to $ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2$ as a *Riemannian metric*. The symmetric matrix associated with the first fundamental form is also denoted by

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix},$$

where $g_{12} = g_{21}$.

(2) As in the previous section, if X is not injective, the first fundamental form I_p is not well-defined. What is well-defined is $I_{(u,v)}$. In some sense this is even worse, since one of the main themes of differential geometry is that the metric properties of a surface (or of a manifold) are captured by a Riemannian metric. Again, we will not worry too much about this, or we will assume X injective.

(3) It can be shown that the element of area dA on a surface X is given by

$$dA = \|X_u \times X_v\|\,du\,dv = \sqrt{EG - F^2}\,du\,dv.$$

We have just discovered that, in contrast to a flat surface, where the inner product is the same at every point, on a curved surface the inner product induced by the Riemannian metric on the tangent space at every point changes as the point moves on the surface. This fundamental idea is at the heart of the definition of an abstract Riemannian manifold. It is also important to observe that the first fundamental form of a surface does **not** characterize the surface.

Example 20.3. It is easy to see that the first fundamental form of a plane and the first fundamental form of a cylinder of revolution defined by

$$X(u,v) = (\cos u, \sin u, v)$$

are identical:

$$(E,F,G) = (1,0,1).$$

Thus $ds^2 = du^2 + dv^2$, which is not surprising.

A more striking example is that of the helicoid and the catenoid.

Example 20.4. The *helicoid* is the surface defined over $\mathbb{R} \times \mathbb{R}$ such that

$$\begin{aligned} x &= u_1 \cos v_1, \\ y &= u_1 \sin v_1, \\ z &= v_1. \end{aligned}$$

This is the surface generated by a line parallel to the xOy plane, touching the z-axis, and also touching a helix of axis Oz. It is easily verified that

$$(E,F,G) = (1,0,u_1^2+1).$$

Figure 20.2 shows a portion of helicoid corresponding to $0 \leq v_1 \leq 2\pi$.

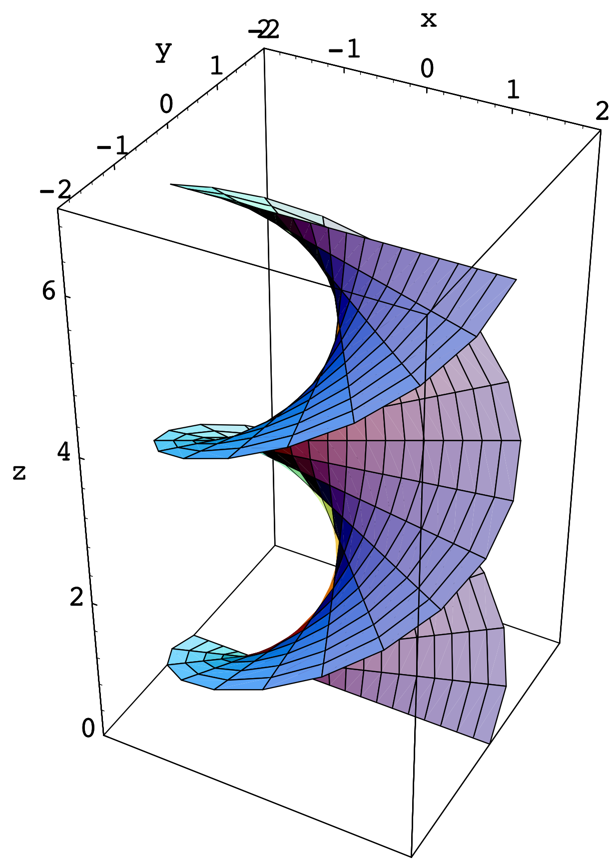

Fig. 20.2 A helicoid.

Example 20.5. The *catenoid* is the surface of revolution defined over $\mathbb{R} \times \mathbb{R}$ such that

$$\begin{aligned} x &= \cosh u_2 \cos v_2, \\ y &= \cosh u_2 \sin v_2, \\ z &= u_2. \end{aligned}$$

It is the surface obtained by rotating a *catenary* around the z-axis. (Recall that the *hyperbolic functions* cosh and sinh are defined by $\cosh u = (e^u + e^{-u})/2$ and $\sinh u = (e^u - e^{-u})/2$. The catenary is the plane curve defined by $y = \cosh x$). It is easily verified that

$$(E, F, G) = (\cosh^2 u_2, 0, \cosh^2 u_2).$$

Figure 20.3 shows a portion of catenoid corresponding to $0 \leq v_2 \leq 2\pi$.

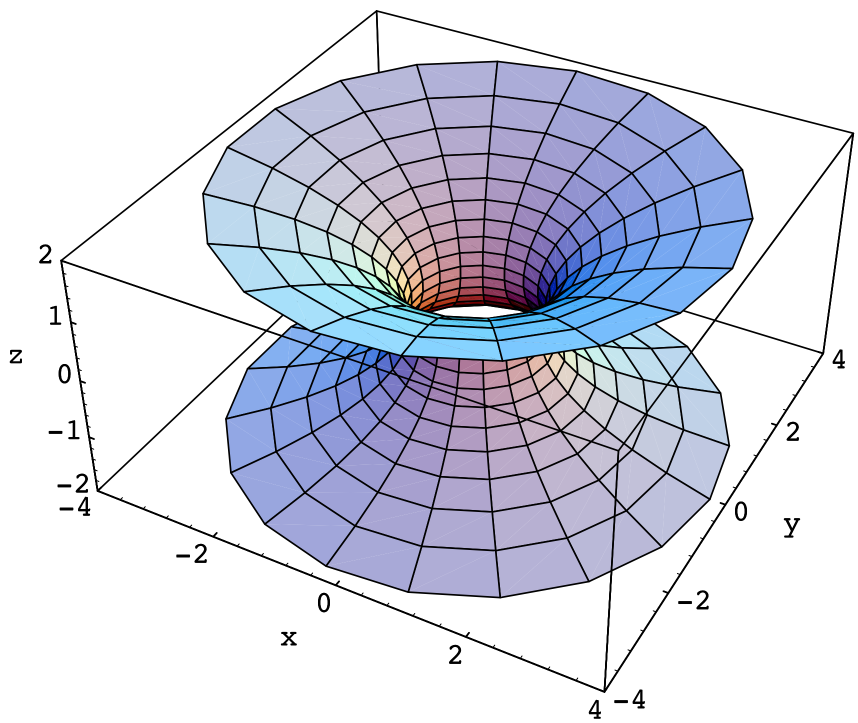

Fig. 20.3 A catenoid.

We can make the change of variables $u_1 = \sinh u_3$, $v_1 = v_3$, which is bijective and whose Jacobian determinant is $\cosh u_3$, which is always positive, obtaining the following parametrization of the helicoid:

$$
\begin{aligned}
x &= \sinh u_3 \cos v_3, \\
y &= \sinh u_3 \sin v_3, \\
z &= v_3.
\end{aligned}
$$

It is easily verified that

$$(E, F, G) = (\cosh^2 u_3, 0, \cosh^2 u_3),$$

showing that the helicoid and the catenoid have the same first fundamental form. What is happening is that the two surfaces are locally isometric (roughly, this means that there is a smooth map between the two surfaces that preserves distances locally). Indeed, if we consider the portions of the two surfaces corresponding to the domain $\mathbb{R}\times\,]0,2\pi[$, it is possible to deform isometrically the portion of helicoid into the portion of catenoid (note that by excluding 0 and 2π, we have made a "slit" in the catenoid (a portion of meridian), and thus we can open up the catenoid and deform it into the helicoid). For more on this, we urge our readers to consult do Carmo [12], Chapter 4, Section 2, pages 218–227.

We will now see how the first fundamental form relates to the curvature of curves on a surface.

20.4 Normal Curvature and the Second Fundamental Form

In this section we take a closer look at the curvature at a point of a curve C on a surface X. Assuming that C is parametrized by arc length, we will see that the vector $X''(s)$ (which is equal to $\kappa\mathbf{n}$, where $\mathbf{n}$ is the principal normal to the curve C at p, and κ is the curvature) can be written as

$$\kappa\mathbf{n} = \kappa_N \mathbf{N} + \kappa_g \mathbf{n_g},$$

where $\mathbf{N}$ is the normal to the surface at p, and $\kappa_g \mathbf{n_g}$ is a tangential component normal to the curve. The component κ_N is called the normal curvature. Computing it will lead to the second fundamental form, another very important quadratic form associated with a surface. The component κ_g is called the geodesic curvature. It turns out that it depends only on the first fundamental form, but computing it is quite complicated, and this will lead to the Christoffel symbols.

Let $f\colon\ ]a,b[\,\to \mathbb{E}^3$ be a curve, where f is at least C^3-continuous, and assume that the curve is parametrized by arc length. We saw in Section 19.6 that if $f'(s) \neq 0$ and $f''(s) \neq 0$ for all $s \in\,]a,b[$ (i.e., f is biregular), we can associate to the point $f(s)$ an orthonormal frame $(\mathbf{t}, \mathbf{n}, \mathbf{b})$ called the Frenet frame, where

$$\begin{aligned}
\mathbf{t} &= f'(s),\\
\mathbf{n} &= \frac{f''(s)}{\|f''(s)\|},\\
\mathbf{b} &= \mathbf{t}\times\mathbf{n}.
\end{aligned}$$

The vector $\mathbf{t}$ is the unit *tangent vector*, the vector $\mathbf{n}$ is called the *principal normal*, and the vector $\mathbf{b}$ is called the *binormal*. Furthermore, the curvature κ at s is $\kappa = \|f''(s)\|$, and thus,

$$f''(s) = \kappa\mathbf{n}.$$

The principal normal $\mathbf{n}$ is contained in the osculating plane at s, which is just the plane spanned by $f'(s)$ and $f''(s)$. Recall that since f is parametrized by arc length, the vector $f'(s)$ is a unit vector, and thus $f'(s)\cdot f'(s) = 1$, and by taking derivatives, we get

$$f'(s)\cdot f''(s) = 0,$$

which shows that $f'(s)$ and $f''(s)$ are linearly independent and orthogonal, provided that $f'(s) \neq 0$ and $f''(s) \neq 0$.

Now, if $C\colon t \mapsto X(u(t), v(t))$ is a curve on a surface X, assuming that C is parametrized by arc length, which implies that

$$(s')^2 = E(u')^2 + 2Fu'v' + G(v')^2 = 1,$$

we have

$$\begin{aligned}
X'(s) &= X_u u' + X_v v',\\
X''(s) &= \kappa\mathbf{n},
\end{aligned}$$

and $\mathbf{t} = X_u u' + X_v v'$ is indeed a unit tangent vector to the curve and to the surface, but $\mathbf{n}$ is the principal normal to the curve, and thus it is **not** necessarily orthogonal to the tangent plane $T_p(X)$ at $p = X(u(t), v(t))$.

Thus, if we intend to study how the curvature κ varies as the curve C passing through p changes, the Frenet frame $(\mathbf{t}, \mathbf{n}, \mathbf{b})$ associated with the curve C is not really adequate, since both $\mathbf{n}$ and $\mathbf{b}$ will vary with C (and $\mathbf{n}$ is undefined when $\kappa = 0$). Thus, it is better to pick a frame associated with the normal to the surface at p, and we pick the frame $(\mathbf{t}, \mathbf{n_g}, \mathbf{N})$ defined as follows.

Definition 20.3. Given a surface X, for any curve $C\colon t \mapsto X(u(t), v(t))$ on X and any point p on X, the orthonormal frame $(\mathbf{t}, \mathbf{n_g}, \mathbf{N})$ is defined such that

$$\begin{aligned}
\mathbf{t} &= X_u u' + X_v v',\\
\mathbf{N} &= \frac{X_u \times X_v}{\|X_u \times X_v\|},\\
\mathbf{n_g} &= \mathbf{N}\times\mathbf{t},
\end{aligned}$$

where $\mathbf{N}$ is the normal vector to the surface X at p. The vector $\mathbf{n_g}$ is called the *geodesic normal vector* (for reasons that will become clear later).

Observe that $\mathbf{n_g}$ is the unit normal vector to the curve C contained in the tangent space $T_p(X)$ at p.

If we use the frame $(\mathbf{t}, \mathbf{n_g}, \mathbf{N})$, we will see shortly that $X''(s) = \kappa\mathbf{n}$ can be written as

$$\kappa\mathbf{n} = \kappa_N\mathbf{N} + \kappa_g\mathbf{n_g}.$$

The component $\kappa_N\mathbf{N}$ is the orthogonal projection of $\kappa\mathbf{n}$ onto the normal direction $\mathbf{N}$, and for this reason κ_N is called the *normal curvature of* C *at* p. The component $\kappa_g\mathbf{n_g}$ is the orthogonal projection of $\kappa\mathbf{n}$ onto the tangent space $T_p(X)$ at p.

We now show how to compute the normal curvature. This will uncover the second fundamental form. Using the abbreviations

$$X_{uu} = \frac{\partial^2 X}{\partial u^2}, \quad X_{uv} = \frac{\partial^2 X}{\partial u \partial v}, \quad X_{vv} = \frac{\partial^2 X}{\partial v^2},$$

since $X' = X_u u' + X_v v'$, using the chain rule we get

$$X'' = X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2 + X_u u'' + X_v v''.$$

In order to decompose $X'' = \kappa\mathbf{n}$ into its normal component (along $\mathbf{N}$) and its tangential component, we use a neat trick suggested by Eugenio Calabi. Recall that

$$(u \times v) \times w = (u \cdot w)v - (w \cdot v)u.$$

Using this identity, we have

$$\begin{aligned}
&(\mathbf{N} \times (X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2) \times \mathbf{N} \\
&\quad = (\mathbf{N} \cdot \mathbf{N})(X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2) \\
&\qquad - (\mathbf{N} \cdot (X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2))\mathbf{N}.
\end{aligned}$$

Since $\mathbf{N}$ is a unit vector, we have $\mathbf{N} \cdot \mathbf{N} = 1$, and consequently, since

$$\kappa\mathbf{n} = X'' = X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2 + X_u u'' + X_v v'',$$

we can write

$$\begin{aligned}
\kappa\mathbf{n} = &(\mathbf{N} \cdot (X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2))\mathbf{N} \\
&+ (\mathbf{N} \times (X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2)) \times \mathbf{N} + X_u u'' + X_v v''.
\end{aligned}$$

Thus, it is clear that the normal component is

$$\kappa_N\mathbf{N} = (\mathbf{N} \cdot (X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2))\mathbf{N},$$

and the normal curvature is given by

$$\kappa_N = \mathbf{N}\cdot(X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2).$$

Letting

$$L = \mathbf{N}\cdot X_{uu}, \quad M = \mathbf{N}\cdot X_{uv}, \quad N = \mathbf{N}\cdot X_{vv},$$

we have

$$\kappa_N = L(u')^2 + 2Mu'v' + N(v')^2.$$

It should be noted that some authors (such as do Carmo) use the notation

$$e = \mathbf{N}\cdot X_{uu}, \quad f = \mathbf{N}\cdot X_{uv}, \quad g = \mathbf{N}\cdot X_{vv}.$$

Recalling that

$$\mathbf{N} = \frac{X_u \times X_v}{\|X_u \times X_v\|},$$

using the Lagrange identity

$$(u\cdot v)^2 + \|u\times v\|^2 = \|u\|^2\|v\|^2,$$

we see that

$$\|X_u \times X_v\| = \sqrt{EG - F^2},$$

and $L = \mathbf{N}\cdot X_{uu}$ can be written as

$$L = \frac{(X_u \times X_v)\cdot X_{uu}}{\sqrt{EG-F^2}} = \frac{(X_u, X_v, X_{uu})}{\sqrt{EG-F^2}},$$

where (X_u, X_v, X_{uu}) is the mixed product, i.e., the determinant of the three vectors (similar expressions are obtained for M and N). Some authors (including Gauss himself and Darboux) use the notation

$$D = (X_u, X_v, X_{uu}), \quad D' = (X_u, X_v, X_{uv}), \quad D'' = (X_u, X_v, X_{vv}),$$

and we also have

$$L = \frac{D}{\sqrt{EG-F^2}}, \quad M = \frac{D'}{\sqrt{EG-F^2}}, \quad N = \frac{D''}{\sqrt{EG-F^2}}.$$

These expressions were used by Gauss to prove his famous *Theorema Egregium*.

Since the quadratic form $(x,y) \mapsto Lx^2 + 2Mxy + Ny^2$ plays a very important role in the theory of surfaces, we introduce the following definition.

Definition 20.4. Given a surface X, for any point $p = X(u,v)$ on X, letting

$$L = \mathbf{N}\cdot X_{uu}, \quad M = \mathbf{N}\cdot X_{uv}, \quad N = \mathbf{N}\cdot X_{vv},$$

where $\mathbf{N}$ is the unit normal at p, the quadratic form $(x,y) \mapsto Lx^2 + 2Mxy + Ny^2$ is called the *second fundamental form of X at p*. It is often denoted by II_p. For a curve C on the surface X (parametrized by arc length), the quantity κ_N given by the formula

$$\kappa_N = L(u')^2 + 2Mu'v' + N(v')^2$$

is called the *normal curvature of C at p*.

The second fundamental form was introduced by Gauss in 1827. Unlike the first fundamental form, the second fundamental form is not necessarily positive or definite. Properties of the surface expressible in terms of the first fundamental form are called *intrinsic properties* of the surface X. Properties of the surface expressible in terms of the second fundamental form are called *extrinsic properties* of the surface X. They have to do with the way the surface is immersed in $\mathbb{E}^3$. As we shall see later, certain notions that appear to be extrinsic turn out to be intrinsic, such as the geodesic curvature and the Gaussian curvature. This is another testimony to the genius of Gauss (and Bonnet, Christoffel, et al.).

Remark: As in the previous section, if X is not injective, the second fundamental form II_p is not well-defined. Again, we will not worry too much about this, or we assume X injective.

It should also be mentioned that the fact that the normal curvature is expressed as

$$\kappa_N = L(u')^2 + 2Mu'v' + N(v')^2$$

has the following immediate corollary, known as *Meusnier's theorem* (1776).

Lemma 20.1. *All curves on a surface X and having the same tangent line at a given point $p \in X$ have the same normal curvature at p.*

In particular, if we consider the curves obtained by intersecting the surface with planes containing the normal at p, curves called *normal sections*, all curves tangent to a normal section at p have the same normal curvature as the normal section. Furthermore, the principal normal of a normal section is collinear with the normal to the surface, and thus $|\kappa| = |\kappa_N|$, where κ is the curvature of the normal section, and κ_N is the normal curvature of the normal section. We will see in a later section how the curvature of normal sections varies.

We obtained the value of the normal curvature κ_N assuming that the curve C is parametrized by arc length, but we can easily give an expression for κ_N for an arbitrary parametrization. Indeed, remember that

$$\left(\frac{ds}{dt}\right)^2 = \|\dot{C}\|^2 = E\,\dot{u}^2 + 2F\,\dot{u}\dot{v} + G\,\dot{v}^2,$$

and by the chain rule

$$u' = \frac{du}{ds} = \frac{du}{dt}\frac{dt}{ds},$$

and since a change of parameter is a diffeomorphism, we get

$$u' = \frac{\dot{u}}{\left(\frac{ds}{dt}\right)},$$

and from
$$\kappa_N = L(u')^2 + 2Mu'v' + N(v')^2,$$
we get
$$\boxed{\kappa_N = \frac{L\dot{u}^2 + 2M\dot{u}\dot{v} + N\dot{v}^2}{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}.}$$

It is remarkable that this expression of the normal curvature uses both the first and the second fundamental forms!

We still need to compute the tangential part X_t'' of X''. We found that the tangential part of X'' is
$$X_t'' = (\mathbf{N} \times (X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2)) \times \mathbf{N} + X_u u'' + X_v v''.$$
This vector is clearly in the tangent space $T_p(X)$ (since the first part is orthogonal to $\mathbf{N}$, which is orthogonal to the tangent space). Furthermore, X'' is orthogonal to X' (since $X' \cdot X' = 1$), and by dotting $X'' = \kappa_N \mathbf{N} + X_t''$ with $\mathbf{t} = X'$, since the component $\kappa_N \mathbf{N} \cdot \mathbf{t}$ is zero, we have $X_t'' \cdot \mathbf{t} = 0$, and thus X_t'' is also orthogonal to $\mathbf{t}$, which means that it is collinear with $\mathbf{n_g} = \mathbf{N} \times \mathbf{t}$. Therefore, we have shown that
$$\kappa \mathbf{n} = \kappa_N \mathbf{N} + \kappa_g \mathbf{n_g},$$
where
$$\kappa_N = L(u')^2 + 2Mu'v' + N(v')^2$$
and
$$\kappa_g \mathbf{n_g} = (\mathbf{N} \times (X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2)) \times \mathbf{N} + X_u u'' + X_v v''.$$
The term $\kappa_g \mathbf{n_g}$ is worth an official definition.

Definition 20.5. Given a surface X, for any curve $C\colon t \mapsto X(u(t), v(t))$ on X and any point p on X, the quantity κ_g appearing in the expression
$$\kappa \mathbf{n} = \kappa_N \mathbf{N} + \kappa_g \mathbf{n_g}$$
giving the acceleration vector of X at p is called the *geodesic curvature of C at p*.

In the next section we give an expression for $\kappa_g \mathbf{n_g}$ in terms of the basis (X_u, X_v).

20.5 Geodesic Curvature and the Christoffel Symbols

We showed that the tangential part of the curvature of a curve C on a surface is of the form $\kappa_g \mathbf{n_g}$. We now show that κ_g can be computed only in terms of the first fundamental form of X, a result first proved by Ossian Bonnet circa 1848. The computation is a bit involved, and it will lead us to the Christoffel symbols, introduced in 1869.

Since $\mathbf{n_g}$ is in the tangent space $T_p(X)$, and since (X_u, X_v) is a basis of $T_p(X)$, we can write

$$\kappa_g \mathbf{n_g} = AX_u + BX_v,$$

for some $A, B \in \mathbb{R}$. However,

$$\kappa \mathbf{n} = \kappa_N \mathbf{N} + \kappa_g \mathbf{n_g},$$

and since $\mathbf{N}$ is normal to the tangent space, $\mathbf{N} \cdot X_u = \mathbf{N} \cdot X_v = 0$, and by dotting

$$\kappa_g \mathbf{n_g} = AX_u + BX_v$$

with X_u and X_v, since $E = X_u \cdot X_u$, $F = X_u \cdot X_v$, and $G = X_v \cdot X_v$, we get the equations

$$\begin{aligned} \kappa \mathbf{n} \cdot X_u &= EA + FB, \\ \kappa \mathbf{n} \cdot X_v &= FA + GB. \end{aligned}$$

On the other hand,

$$\kappa \mathbf{n} = X'' = X_u u'' + X_v v'' + X_{uu}(u')^2 + 2X_{uv} u'v' + X_{vv}(v')^2.$$

Dotting with X_u and X_v, we get

$$\begin{aligned} \kappa \mathbf{n} \cdot X_u &= Eu'' + Fv'' + (X_{uu} \cdot X_u)(u')^2 + 2(X_{uv} \cdot X_u)u'v' + (X_{vv} \cdot X_u)(v')^2, \\ \kappa \mathbf{n} \cdot X_v &= Fu'' + Gv'' + (X_{uu} \cdot X_v)(u')^2 + 2(X_{uv} \cdot X_v)u'v' + (X_{vv} \cdot X_v)(v')^2. \end{aligned}$$

At this point it is useful to introduce the *Christoffel symbols (of the first kind)* $[\alpha\beta;\gamma]$, defined such that

$$[\alpha\beta;\gamma] = X_{\alpha\beta} \cdot X_\gamma,$$

where $\alpha, \beta, \gamma \in \{u, v\}$. It is also more convenient to let $u = u_1$ and $v = u_2$, and to denote $[u_\alpha v_\beta; u_\gamma]$ by $[\alpha\beta;\gamma]$. Doing so, and remembering that

$$\begin{aligned} \kappa \mathbf{n} \cdot X_u &= EA + FB, \\ \kappa \mathbf{n} \cdot X_v &= FA + GB, \end{aligned}$$

we have the following equation:

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} u_1'' \\ u_2'' \end{pmatrix} + \sum_{\substack{\alpha=1,2 \\ \beta=1,2}} \begin{pmatrix} [\alpha\beta;1]\, u'_\alpha u'_\beta \\ [\alpha\beta;2]\, u'_\alpha u'_\beta \end{pmatrix}.$$

However, since the first fundamental form is positive definite, $EG - F^2 > 0$, and we have

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} = (EG - F^2)^{-1} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix}.$$

Thus, we get

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} u_1'' \\ u_2'' \end{pmatrix} + \sum_{\substack{\alpha=1,2 \\ \beta=1,2}} (EG-F^2)^{-1} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix} \begin{pmatrix} [\alpha\,\beta;\,1]\, u_\alpha' u_\beta' \\ [\alpha\,\beta;\,2]\, u_\alpha' u_\beta' \end{pmatrix}.$$

It is natural to introduce the *Christoffel symbols (of the second kind)* $\Gamma_{i\,j}^k$, defined such that

$$\begin{pmatrix} \Gamma_{i\,j}^1 \\ \Gamma_{i\,j}^2 \end{pmatrix} = (EG-F^2)^{-1} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix} \begin{pmatrix} [i\,j;\,1] \\ [i\,j;\,2] \end{pmatrix}.$$

Finally, we get

$$A = u_1'' + \sum_{\substack{i=1,2 \\ j=1,2}} \Gamma_{i\,j}^1 u_i' u_j',$$
$$B = u_2'' + \sum_{\substack{i=1,2 \\ j=1,2}} \Gamma_{i\,j}^2 u_i' u_j',$$

and

$$\kappa_g \mathbf{n_g} = \Bigg(u_1'' + \sum_{\substack{i=1,2 \\ j=1,2}} \Gamma_{i\,j}^1 u_i' u_j' \Bigg) X_u + \Bigg(u_2'' + \sum_{\substack{i=1,2 \\ j=1,2}} \Gamma_{i\,j}^2 u_i' u_j' \Bigg) X_v.$$

We summarize all the above in the following lemma.

Lemma 20.2. *Given a surface X and a curve C on X, for any point p on C, the tangential part of the curvature at p is given by*

$$\kappa_g \mathbf{n_g} = \Bigg(u_1'' + \sum_{\substack{i=1,2 \\ j=1,2}} \Gamma_{i\,j}^1 u_i' u_j' \Bigg) X_u + \Bigg(u_2'' + \sum_{\substack{i=1,2 \\ j=1,2}} \Gamma_{i\,j}^2 u_i' u_j' \Bigg) X_v,$$

where the Christoffel symbols $\Gamma_{i\,j}^k$ are defined such that

$$\begin{pmatrix} \Gamma_{i\,j}^1 \\ \Gamma_{i\,j}^2 \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} [i\,j;\,1] \\ [i\,j;\,2] \end{pmatrix},$$

and the Christoffel symbols $[i\,j;\,k]$ are defined such that

$$[i\,j;\,k] = X_{ij} \cdot X_k.$$

Looking at the formulae

$$[\alpha\,\beta;\,\gamma] = X_{\alpha\beta} \cdot X_\gamma$$

for the Christoffel symbols $[\alpha\,\beta;\,\gamma]$, it does not seem that these symbols depend only on the first fundamental form, but in fact, they do! Firstly, note that

$$[\alpha\,\beta;\,\gamma] = [\beta\,\alpha;\,\gamma].$$

Next, observe that

$$\begin{aligned}
X_{uu} \cdot X_u &= \frac{1}{2}\frac{\partial(X_u \cdot X_u)}{\partial u} = \frac{1}{2}E_u,\\
X_{uv} \cdot X_u &= \frac{1}{2}\frac{\partial(X_u \cdot X_u)}{\partial v} = \frac{1}{2}E_v,\\
X_{uv} \cdot X_v &= \frac{1}{2}\frac{\partial(X_v \cdot X_v)}{\partial u} = \frac{1}{2}G_u,\\
X_{vv} \cdot X_v &= \frac{1}{2}\frac{\partial(X_v \cdot X_v)}{\partial v} = \frac{1}{2}G_v,
\end{aligned}$$

and since

$$(X_u \cdot X_v)_v = X_{uv} \cdot X_v + X_u \cdot X_{vv}$$

and

$$X_{uv} \cdot X_v = \frac{1}{2}G_u,$$

we get

$$F_v = \frac{1}{2}G_u + X_u \cdot X_{vv},$$

and thus

$$X_{vv} \cdot X_u = F_v - \frac{1}{2}G_u.$$

Similarly, we get

$$X_{uu} \cdot X_v = F_u - \frac{1}{2}E_v.$$

In summary, we have the following formulae showing that the Christoffel symbols depend only on the first fundamental form:

$$\begin{aligned}
&[1\,1;1] = \frac{1}{2}E_u, \quad [1\,1;2] = F_u - \frac{1}{2}E_v,\\
&[1\,2;1] = \frac{1}{2}E_v, \quad [1\,2;2] = \frac{1}{2}G_u,\\
&[2\,1;1] = \frac{1}{2}E_v, \quad [2\,1;2] = \frac{1}{2}G_u,\\
&[2\,2;1] = F_v - \frac{1}{2}G_u, \quad [2\,2;2] = \frac{1}{2}G_v.
\end{aligned}$$

Another way to compute the Christoffel symbols $[\alpha\,\beta;\gamma]$, is to proceed as follows. For this computation it is more convenient to assume that $u = u_1$ and $v = u_2$, and that the first fundamental form is expressed by the matrix

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix},$$

where $g_{\alpha\beta} = X_\alpha \cdot X_\beta$. Let

$$g_{\alpha\beta\,|\,\gamma} = \frac{\partial g_{\alpha\beta}}{\partial u_\gamma}.$$

Then, we have

$$g_{\alpha\beta\,|\,\gamma} = \frac{\partial g_{\alpha\beta}}{\partial u_\gamma} = X_{\alpha\gamma}\cdot X_\beta + X_\alpha \cdot X_{\beta\gamma} = [\alpha\,\gamma;\,\beta] + [\beta\,\gamma;\,\alpha].$$

From this, we also have

$$g_{\beta\gamma\,|\,\alpha} = [\alpha\,\beta;\,\gamma] + [\alpha\,\gamma;\,\beta]$$

and

$$g_{\alpha\gamma\,|\,\beta} = [\alpha\,\beta;\,\gamma] + [\beta\,\gamma;\,\alpha].$$

From all this we get

$$2[\alpha\,\beta;\,\gamma] = g_{\alpha\gamma\,|\,\beta} + g_{\beta\gamma\,|\,\alpha} - g_{\alpha\beta\,|\,\gamma}.$$

As before, the Christoffel symbols $[\alpha\,\beta;\,\gamma]$ and $\Gamma^\gamma_{\alpha\beta}$ are related via the Riemannian metric by the equations

$$\Gamma^\gamma_{\alpha\beta} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}^{-1} [\alpha\,\beta;\,\gamma].$$

This seemingly bizarre approach has the advantage of generalizing to Riemannian manifolds. In the next section we study the variation of the normal curvature.

20.6 Principal Curvatures, Gaussian Curvature, Mean Curvature

We will now study how the normal curvature at a point varies when a unit tangent vector varies. In general, we will see that the normal curvature has a maximum value κ_1 and a minimum value κ_2, and that the corresponding directions are orthogonal. This was shown by Euler in 1760. The quantity $K = \kappa_1\kappa_2$, called the Gaussian curvature, and the quantity $H = (\kappa_1 + \kappa_2)/2$, called the mean curvature, play a very important role in the theory of surfaces. We will compute H and K in terms of the first and the second fundamental forms. We also classify points on a surface according to the value and sign of the Gaussian curvature.

Recall that given a surface X and some point p on X, the vectors X_u, X_v form a basis of the tangent space $T_p(X)$. Given a unit vector $\mathbf{t} = X_u x + X_v y$, the normal curvature is given by

$$\kappa_N(\mathbf{t}) = Lx^2 + 2Mxy + Ny^2,$$

since $Ex^2 + 2Fxy + Gy^2 = 1$. Usually, (X_u, X_v) is not an orthonormal frame, and it is useful to replace the frame (X_u, X_v) with an orthonormal frame. One verifies easily that the frame (e_1, e_2) defined such that

$$e_1 = \frac{X_u}{\sqrt{E}}, \quad e_2 = \frac{EX_v - FX_u}{\sqrt{E(EG - F^2)}}$$

is indeed an orthonormal frame. With respect to this frame, every unit vector can be written as $\mathbf{t} = \cos\theta e_1 + \sin\theta e_2$, and expressing (e_1, e_2) in terms of X_u and X_v, we have

$$\mathbf{t} = \left(\frac{w\cos\theta - F\sin\theta}{w\sqrt{E}}\right) X_u + \frac{\sqrt{E}\sin\theta}{w} X_v,$$

where $w = \sqrt{EG - F^2}$. We can now compute $\kappa_N(\mathbf{t})$, and we get

$$\kappa_N(\mathbf{t}) = L\left(\frac{w\cos\theta - F\sin\theta}{w\sqrt{E}}\right)^2 + 2M\left(\frac{(w\cos\theta - F\sin\theta)\sin\theta}{w^2}\right) + N\frac{E\sin^2\theta}{w^2}.$$

We leave as an exercise to show that the above expression can be written as

$$\kappa_N(\mathbf{t}) = H + A\cos 2\theta + B\sin 2\theta,$$

where

$$\begin{aligned} H &= \frac{GL - 2FM + EN}{2(EG - F^2)}, \\ A &= \frac{L(EG - 2F^2) + 2EFM - E^2N}{2E(EG - F^2)}, \\ B &= \frac{EM - FL}{E\sqrt{EG - F^2}}. \end{aligned}$$

Letting $C = \sqrt{A^2 + B^2}$, unless $A = B = 0$, the function

$$f(\theta) = H + A\cos 2\theta + B\sin 2\theta$$

has a maximum $\kappa_1 = H + C$ for the angles θ_0 and $\theta_0 + \pi$, and a minimum $\kappa_2 = H - C$ for the angles $\theta_0 + \pi/2$ and $\theta_0 + 3\pi/2$, where $\cos 2\theta_0 = A/C$ and $\sin 2\theta_0 = B/C$. The curvatures κ_1 and κ_2 play a major role in surface theory.

Definition 20.6. Given a surface X, for any point p on X, letting A, B, H be defined as above, and $C = \sqrt{A^2 + B^2}$, unless $A = B = 0$, the normal curvature κ_N at p takes a maximum value κ_1 and and a minimum value κ_2, called *principal curvatures at p*, where $\kappa_1 = H + C$ and $\kappa_2 = H - C$. The directions of the corresponding unit vectors are called the *principal directions at p*. The average $H = \kappa_1 + \kappa_2/2$ of the principal curvatures is called the *mean curvature*, and the product $K = \kappa_1\kappa_2$ of the principal curvatures is called the *total curvature*, or *Gaussian curvature*.

Observe that the principal directions θ_0 and $\theta_0 + \pi/2$ corresponding to κ_1 and κ_2 are orthogonal. Note that

$$K = \kappa_1 \kappa_2 = (H - C)(H + C) = H^2 - C^2 = H^2 - (A^2 + B^2).$$

We leave as an exercise to verify that

$$A^2 + B^2 = \frac{G^2L^2 - 4FGLM + 4EGM^2 + 4F^2LN - 2EGLN - 4EFMN + E^2N^2}{4(EG - F^2)^2}$$

and

$$H^2 = \frac{G^2L^2 - 4FGLM + 4F^2M^2 + 2EGLN - 4EFMN + E^2N^2}{4(EG - F^2)^2}.$$

From this we get

$$H^2 - A^2 - B^2 = \frac{LN - M^2}{EG - F^2}.$$

In summary, we have the following (famous) formulae for the mean curvature and the Gaussian curvature:

$$\boxed{\begin{aligned} H &= \frac{GL - 2FM + EN}{2(EG - F^2)}, \\ K &= \frac{LN - M^2}{EG - F^2}. \end{aligned}}$$

We have shown that the normal curvature κ_N can be expressed as

$$\kappa_N(\theta) = H + A\cos 2\theta + B\sin 2\theta$$

over the orthonormal frame (e_1, e_2). We also have shown that the angle θ_0 such that $\cos 2\theta_0 = A/C$ and $\sin 2\theta_0 = B/C$ plays a special role. Indeed, it determines one of the principal directions. If we rotate the basis (e_1, e_2) and pick a frame (f_1, f_2) corresponding to the principal directions, we obtain a particularly nice formula for κ_N. Indeed, since $A = C\cos\ 2\theta_0$ and $B = C\sin 2\theta_0$, letting $\varphi = \theta - \theta_0$, we can write

$$\begin{aligned}
\kappa_N(\theta) &= H + A\cos 2\theta + B\sin 2\theta \\
&= H + C(\cos 2\theta_0 \cos 2\theta + \sin 2\theta_0 \sin 2\theta) \\
&= H + C(\cos 2(\theta - \theta_0)) \\
&= H + C(\cos^2(\theta - \theta_0) - \sin^2(\theta - \theta_0)) \\
&= H(\cos^2(\theta - \theta_0) + \sin^2(\theta - \theta_0)) + C(\cos^2(\theta - \theta_0) - \sin^2(\theta - \theta_0)) \\
&= (H + C)\cos^2(\theta - \theta_0) + (H - C)\sin^2(\theta - \theta_0) \\
&= \kappa_1 \cos^2\varphi + \kappa_2 \sin^2\varphi.
\end{aligned}$$

Thus, for any unit vector $\mathbf{t}$ expressed as

$$\mathbf{t} = \cos\varphi f_1 + \sin\varphi f_2$$

with respect to an orthonormal frame corresponding to the principal directions, the normal curvature $\kappa_N(\varphi)$ is given by *Euler's formula* (1760)

$$\boxed{\kappa_N(\varphi) = \kappa_1 \cos^2\varphi + \kappa_2 \sin^2\varphi.}$$

Recalling that $EG - F^2$ is always strictly positive, we can classify the points on the surface depending on the value of the Gaussian curvature K and on the values of the principal curvatures κ_1 and κ_2 (or H).

Definition 20.7. Given a surface X, a point p on X belongs to one of the following categories:

(1) *Elliptic* if $LN - M^2 > 0$, or equivalently $K > 0$.
(2) *Hyperbolic* if $LN - M^2 < 0$, or equivalently $K < 0$.
(3) *Parabolic* if $LN - M^2 = 0$ and $L^2 + M^2 + N^2 > 0$, or equivalently $K = \kappa_1\kappa_2 = 0$ but either $\kappa_1 \neq 0$ or $\kappa_2 \neq 0$.
(4) *Planar* if $L = M = N = 0$, or equivalently $\kappa_1 = \kappa_2 = 0$.

Furthermore, a point p is an *umbilical point* (or *umbilic*) if $K > 0$ and $\kappa_1 = \kappa_2$.

Note that some authors allow a planar point to be an umbilical point, but we do not. At an elliptic point, both principal curvatures are nonnull and have the same sign. For example, most points on an ellipsoid are elliptic.

At a hyperbolic point, the principal curvatures have opposite signs. For example, all points on the catenoid are hyperbolic.

At a parabolic point, one of the two principal curvatures is zero, but not both. This is equivalent to $K = 0$ and $H \neq 0$. Points on a cylinder are parabolic.

At a planar point, $\kappa_1 = \kappa_2 = 0$. This is equivalent to $K = H = 0$. Points on a plane are all planar points!

Example 20.6. On a monkey saddle, there is a planar point, as shown in Figure 20.4. The principal directions at that point are undefined.

For an umbilical point we have $\kappa_1 = \kappa_2 \neq 0$. This can happen only when $H - C = H + C$, which implies that $C = 0$, and since $C = \sqrt{A^2 + B^2}$, we have $A = B = 0$. Thus, for an umbilical point, $K = H^2$. In this case the function κ_N is constant, and the principal directions are undefined. All points on a sphere are umbilics. A general ellipsoid (a, b, c pairwise distinct) has four umbilics.

It can be shown that a connected surface consisting only of umbilical points is contained in a sphere (see do Carmo [12], Section 3.2, or Gray [23], Section 28.2). It can also be shown that a connected surface consisting only of planar points is contained in a plane. A surface can contain at the same time elliptic points, parabolic points, and hyperbolic points. This is the case of a torus.

Example 20.7. The parabolic points are on two circles also contained in two tangent planes to the torus (the two horizontal planes touching the top and the bottom of the torus, as shown in Figure 20.5). The elliptic points are on the outside part of the torus (with normal facing outward), delimited by the two circles of parabolic

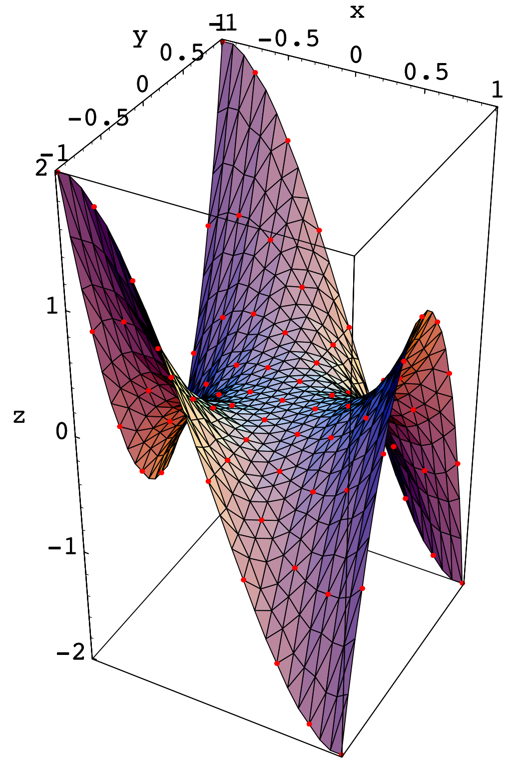

Fig. 20.4 A monkey saddle.

points. The hyperbolic points are on the inside part of the torus (with normal facing inward).

The normal curvature $\kappa_N(X_u x + X_v y) = Lx^2 + 2Mxy + Ny^2$ will vanish for some tangent vector $(x, y) \neq (0, 0)$ iff $M^2 - LN \geq 0$. Since

$$K = \frac{LN - M^2}{EG - F^2},$$

this can happen only if $K \leq 0$. If $L = N = 0$, then there are two directions corresponding to X_u and X_v for which the normal curvature is zero. If $L \neq 0$ or $N \neq 0$, say $L \neq 0$ (the other case being similar), then the equation

$$L\left(\frac{x}{y}\right)^2 + 2M\frac{x}{y} + N = 0$$

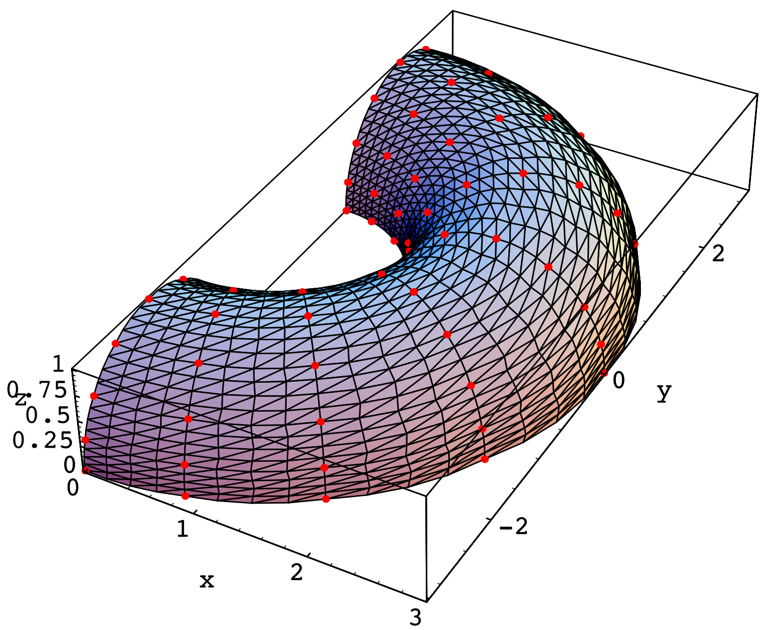

Fig. 20.5 Portion of torus.

has two distinct roots iff $K < 0$. The directions corresponding to the vectors $X_u x + X_v y$ associated with these roots are called the *asymptotic directions at* p. These are the directions for which the normal curvature is null at p.

There are surfaces of constant Gaussian curvature. For example, a cylinder or a cone is a surface of Gaussian curvature $K = 0$. A sphere of radius R has positive constant Gaussian curvature $K = 1/R^2$. Perhaps surprisingly, there are other surfaces of constant positive curvature besides the sphere. There are surfaces of constant negative curvature, say $K = -1$.

Example 20.8. A famous surfaces of constant negative curvature is the *pseudo-sphere*, also known as *Beltrami's pseudosphere*. This is the surface of revolution obtained by rotating a curve known as a *tractrix* around its asymptote. One possible parametrization is given by

$$
\begin{aligned}
x &= \frac{2\cos v}{e^u + e^{-u}},\\
y &= \frac{2\sin v}{e^u + e^{-u}},\\
z &= u - \frac{e^u - e^{-u}}{e^u + e^{-u}},
\end{aligned}
$$

over $]0, 2\pi[\,\times\mathbb{R}$. The pseudosphere has a circle of singular points (for $u = 0$). Figure 20.6 shows a portion of pseudosphere.

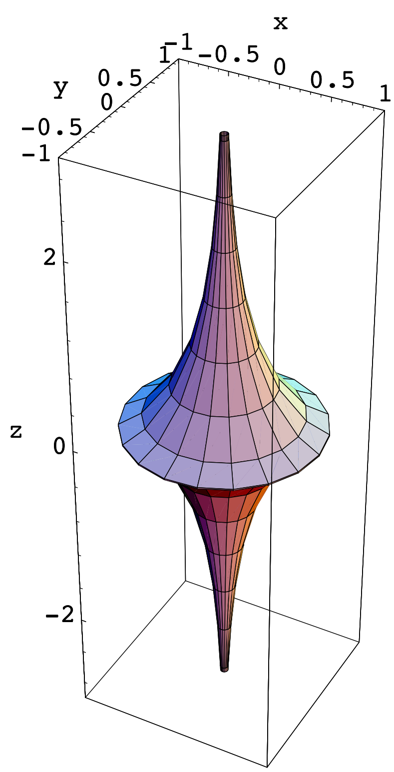

Fig. 20.6 A pseudosphere.

Again, perhaps surprisingly, there are other surfaces of constant negative curvature.

The Gaussian curvature at a point (x, y, x) of an ellipsoid of equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

has the beautiful expression

$$K = \frac{p^4}{a^2b^2c^2},$$

where p is the distance from the origin $(0,0,0)$ to the tangent plane at the point (x,y,z).

There are also surfaces for which $H = 0$. Such surfaces are called *minimal surfaces*, and they show up in physics quite a bit. It can be verified that both the helicoid and the catenoid are minimal surfaces. The Enneper surface is also a minimal surface (see Example 20.9).

We will see shortly how the classification of points on a surface can be explained in terms of the Dupin indicatrix. The idea is to dip the surface in water, and to watch the shorelines formed in the water by the surface in a small region around a chosen point, as we move the surface up and down very gently. But first, we introduce the Gauss map, i.e., we study the variations of the normal $\mathbf{N}_p$ as the point p varies on the surface.

20.7 The Gauss Map and Its Derivative $d\mathbf{N}$

Given a surface $X\colon \Omega \to \mathbb{E}^3$ and any point $p = X(u,v)$ on X, we have defined the normal $\mathbf{N}_p$ at p (or really $\mathbf{N}_{(u,v)}$ at (u,v)) as the unit vector

$$\mathbf{N}_p = \frac{X_u \times X_v}{\|X_u \times X_v\|}.$$

Gauss realized that the assignment $p \mapsto \mathbf{N}_p$ of the unit normal $\mathbf{N}_p$ to the point p on the surface X could be viewed as a map from the trace of the surface X to the unit sphere S^2. If $\mathbf{N}_p$ is a unit vector of coordinates (x,y,z), we have $x^2+y^2+z^2 = 1$, and $\mathbf{N}_p$ corresponds to the point $N(p) = (x,y,z)$ on the unit sphere. This is the so-called *Gauss map of* X, denoted by $\mathbf{N}\colon X \to S^2$.

The derivative $d\mathbf{N}_p$ of the Gauss map at p measures the variation of the normal near p, i.e., how the surface "curves" near p. The Jacobian matrix of $d\mathbf{N}_p$ in the basis (X_u, X_v) can be expressed simply in terms of the matrices associated with the first and the second fundamental forms (which are quadratic forms). Furthermore, the eigenvalues of $d\mathbf{N}_p$ are precisely $-\kappa_1$ and $-\kappa_2$, where κ_1 and κ_2 are the principal curvatures at p, and the eigenvectors define the principal directions (when they are well-defined). In view of the negative sign in $-\kappa_1$ and $-\kappa_2$, it is desirable to consider the linear map $\mathscr{S}_p = -d\mathbf{N}_p$, often called the *shape operator*. Then it is easily shown that the second fundamental form $\mathrm{II}_p(\mathbf{t})$ can be expressed as

$$\mathrm{II}_p(\mathbf{t}) = \langle \mathscr{S}_p(\mathbf{t}), \mathbf{t}\rangle_p,$$

where $\langle -,-\rangle$ is the inner product associated with the first fundamental form. Thus, the Gaussian curvature is equal to the determinant of $\mathscr{S}_p$, and also to the determinant of $d\mathbf{N}_p$, since $(-\kappa_1)(-\kappa_2) = \kappa_1\kappa_2$. We will see in a later section that the Gaussian curvature actually depends only of the first fundamental form, which is far from obvious right now!

Actually, if X is not injective, there are problems, because the assignment $p \mapsto \mathbf{N}_p$ could be multivalued, since there could be several different normals. We can either assume that X is injective, or consider the map from Ω to S^2 defined such that

$$(u,v) \mapsto \mathbf{N}_{(u,v)}.$$

Then we have a map from Ω to S^2, where (u,v) is mapped to the point $N(u,v)$ on S^2 associated with $\mathbf{N}_{(u,v)}$. This map is denoted by $\mathbf{N}\colon \Omega \to S^2$.

It is interesting to study the derivative $d\mathbf{N}$ of the Gauss map $\mathbf{N}\colon \Omega \to S^2$ (or $\mathbf{N}\colon X \to S^2$). As we shall see, the second fundamental form can be defined in terms of $d\mathbf{N}$. For every $(u,v) \in \Omega$, the map $d\mathbf{N}_{(u,v)}$ is a linear map $d\mathbf{N}_{(u,v)}\colon \mathbb{R}^2 \to \mathbb{R}^2$. It can be viewed as a linear map from the tangent space $T_{(u,v)}(X)$ at $X(u,v)$ (which is isomorphic to $\mathbb{R}^2$) to the tangent space to the sphere at $N(u,v)$ (also isomorphic to $\mathbb{R}^2$). Recall that $d\mathbf{N}_{(u,v)}$ is defined as follows: For every $(x,y) \in \mathbb{R}^2$,

$$d\mathbf{N}_{(u,v)}(x,y) = \mathbf{N}_u x + \mathbf{N}_v y.$$

Thus, we need to compute $\mathbf{N}_u$ and $\mathbf{N}_v$. Since $\mathbf{N}$ is a unit vector, $\mathbf{N} \cdot \mathbf{N} = 1$, and by taking derivatives, we have $\mathbf{N}_u \cdot \mathbf{N} = 0$ and $\mathbf{N}_v \cdot \mathbf{N} = 0$. Consequently, $\mathbf{N}_u$ and $\mathbf{N}_v$ are in the tangent space at (u,v), and we can write

$$\begin{aligned}\mathbf{N}_u &= aX_u + cX_v,\\ \mathbf{N}_v &= bX_u + dX_v.\end{aligned}$$

The lemma below shows how to compute a,b,c,d in terms of the first and the second fundamental forms.

Lemma 20.3. *Given a surface X, for any point $p = X(u,v)$ on X, the derivative $d\mathbf{N}_{(u,v)}$ of the Gauss map expressed in the basis (X_u, X_v) is given by the equation*

$$d\mathbf{N}_{(u,v)} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

where the Jacobian matrix $J(d\mathbf{N}_{(u,v)})$ of $d\mathbf{N}_{(u,v)}$ is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = -\begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} L & M \\ M & N \end{pmatrix},$$

that is,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{EG - F^2} \begin{pmatrix} MF - LG & NF - MG \\ LF - ME & MF - NE \end{pmatrix}.$$

Proof. By dotting the equations

$$\mathbf{N}_u = aX_u + cX_v,$$
$$\mathbf{N}_v = bX_u + dX_v,$$

with X_u and X_v, we get

$$\begin{aligned}\mathbf{N}_u \cdot X_u &= aE + cF,\\ \mathbf{N}_u \cdot X_v &= aF + cG,\\ \mathbf{N}_v \cdot X_u &= bE + dF,\\ \mathbf{N}_v \cdot X_v &= bF + dG.\end{aligned}$$

We can compute $\mathbf{N}_u \cdot X_u$, $\mathbf{N}_u \cdot X_v$, $\mathbf{N}_v \cdot X_u$, and $\mathbf{N}_v \cdot X_v$, using the fact that $\mathbf{N} \cdot X_u = \mathbf{N} \cdot X_v = 0$. By taking derivatives, we get

$$\begin{aligned}\mathbf{N} \cdot X_{uu} + \mathbf{N}_u \cdot X_u &= 0,\\ \mathbf{N} \cdot X_{uv} + \mathbf{N}_v \cdot X_u &= 0,\\ \mathbf{N} \cdot X_{vu} + \mathbf{N}_u \cdot X_v &= 0,\\ \mathbf{N} \cdot X_{vv} + \mathbf{N}_v \cdot X_v &= 0.\end{aligned}$$

Thus, we have

$$\begin{aligned}\mathbf{N}_u \cdot X_u &= -L,\\ \mathbf{N}_u \cdot X_v &= -M,\\ \mathbf{N}_v \cdot X_u &= -M,\\ \mathbf{N}_v \cdot X_v &= -N,\end{aligned}$$

and together with the previous equations, we get

$$\begin{aligned}-L &= aE + cF,\\ -M &= aF + cG,\\ -M &= bE + dF,\\ -N &= bF + dG.\end{aligned}$$

This system can be written in matrix form as

$$-\begin{pmatrix} L & M \\ M & N \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

Therefore, we have

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} = -\begin{pmatrix} L & M \\ M & N \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1},$$

which yields

$$\begin{pmatrix} \mathbf{N}_u \\ \mathbf{N}_v \end{pmatrix} = -\begin{pmatrix} L & M \\ M & N \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} X_u \\ X_v \end{pmatrix}.$$

However, we have

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} = \frac{1}{EG-F^2}\begin{pmatrix} G & -F \\ -F & E \end{pmatrix},$$

and thus

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \frac{-1}{EG-F^2}\begin{pmatrix} L & M \\ M & N \end{pmatrix}\begin{pmatrix} G & -F \\ -F & E \end{pmatrix},$$

that is,

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \frac{1}{EG-F^2}\begin{pmatrix} MF-LG & LF-ME \\ NF-MG & MF-NE \end{pmatrix}.$$

We shall now see that the Jacobian matrix $J(d\mathbf{N}_{(u,v)})$ of the linear map $d\mathbf{N}_{(u,v)}$ expressed in the basis (X_u, X_v) is the transpose of the above matrix. Indeed, as we saw earlier,

$$d\mathbf{N}_{(u,v)}(x,y) = \mathbf{N}_u x + \mathbf{N}_v y,$$

and using the expressions for $\mathbf{N}_u$ and $\mathbf{N}_v$, we get

$$d\mathbf{N}_{(u,v)}(x,y) = (aX_u + cX_v)x + (bX_u + dX_v)y = (ax+by)X_u + (cx+dy)X_v,$$

and thus

$$d\mathbf{N}_{(u,v)}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

and since $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the transpose of $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$, we get

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = -\begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1}\begin{pmatrix} L & M \\ M & N \end{pmatrix},$$

that is,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{EG-F^2}\begin{pmatrix} MF-LG & NF-MG \\ LF-ME & MF-NE \end{pmatrix}.$$

This concludes the proof. $\square$

The equations

$$J(d\mathbf{N}_{(u,v)}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{EG-F^2}\begin{pmatrix} MF-LG & NF-MG \\ LF-ME & MF-NE \end{pmatrix}$$

are known as the *Weingarten equations* (in matrix form). If we recall from Section 20.6 the expressions for the Gaussian curvature and for the mean curvature

$$H = \frac{GL - 2FM + EN}{2(EG - F^2)},$$
$$K = \frac{LN - M^2}{EG - F^2},$$

we note that the trace $a+d$ of the Jacobian matrix $J(d\mathbf{N}_{(u,v)})$ of $d\mathbf{N}_{(u,v)}$ is $-2H$, and that its determinant is precisely K. This is recorded in the following lemma, which also shows that the eigenvectors of $J(d\mathbf{N}_{(u,v)})$ correspond to the principal directions.

Lemma 20.4. *Given a surface X, for any point $p = X(u,v)$ on X, the eigenvalues of the Jacobian matrix $J(d\mathbf{N}_{(u,v)})$ of the derivative $d\mathbf{N}_{(u,v)}$ of the Gauss map are $-\kappa_1, -\kappa_2$, where κ_1 and κ_2 are the principal curvatures at p, and the eigenvectors of $d\mathbf{N}_{(u,v)}$ correspond to the principal directions (when they are defined). The Gaussian curvature K is the determinant of the Jacobian matrix of $d\mathbf{N}_{(u,v)}$, and the mean curvature H is equal to $-\frac{1}{2}\mathrm{tr}(J(d\mathbf{N}_{(u,v)}))$.*

Proof. We have just observed that the trace $a+d$ of the matrix

$$J(d\mathbf{N}_{(u,v)}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{EG - F^2} \begin{pmatrix} MF - LG & NF - MG \\ LF - ME & MF - NE \end{pmatrix}$$

is $-2H$, and that its determinant is precisely K. However, the characteristic equation of the above matrix is

$$x^2 - \mathrm{tr}(J(d\mathbf{N}_{(u,v)}))x + \det(J(d\mathbf{N}_{(u,v)})) = 0,$$

which is just

$$x^2 + 2Hx + K = 0.$$

Since $\kappa_1\kappa_2 = K$ and $\kappa_1 + \kappa_2 = 2H$, κ_1 and κ_2 are the roots of the equation

$$x^2 - 2Hx + K = 0.$$

This shows that the eigenvalues of $J(d\mathbf{N}_{(u,v)})$, which are the roots of the equation

$$x^2 + 2Hx + K = 0,$$

are indeed $-\kappa_1$ and $-\kappa_2$.

Recall that κ_1 and κ_2 are the maximum and minimum values of the normal curvature, which is given by the equation

$$\kappa_N(x,y) = \frac{Lx^2 + 2Mxy + Ny^2}{Ex^2 + 2Fxy + Gy^2}.$$

Thus, the partial derivatives $\partial\kappa_N(u',v')/\partial x$ and $\partial\kappa_N(u',v')/\partial y$ of the above function must be zero for the principal directions (u',v') associated with κ_1 and κ_2. It is easy to see that this yields the equations

$$\begin{aligned}(L-\kappa E)u' + (M-\kappa F)v' &= 0,\\ (M-\kappa F)u' + (N-\kappa G)v' &= 0,\end{aligned}$$

where κ is either κ_1 or κ_2. On the other hand, the eigenvectors of $J(d\mathbf{N}_{(u,v)})$ also satisfy the equation

$$J(d\mathbf{N}_{(u,v)})\begin{pmatrix}u'\\ v'\end{pmatrix} = -\kappa\begin{pmatrix}u'\\ v'\end{pmatrix},$$

that is

$$\begin{aligned}\frac{MF-LG}{EG-F^2}u' + \frac{NF-MG}{EG-F^2}v' &= -\kappa u',\\ \frac{LF-ME}{EG-F^2}u' + \frac{MF-NE}{EG-F^2}v' &= -\kappa v',\end{aligned}$$

where $\kappa = \kappa_1$ or $\kappa = \kappa_2$. From the equations

$$\begin{aligned}(L-\kappa E)u' + (M-\kappa F)v' &= 0,\\ (M-\kappa F)u' + (N-\kappa G)v' &= 0,\end{aligned}$$

we get

$$\begin{aligned}Lu' + Mv' &= \kappa(Eu' + Fv'),\\ Mu' + Nv' &= \kappa(Fu' + Gv'),\end{aligned}$$

which reads in matrix form as

$$\begin{pmatrix}L & M\\ M & N\end{pmatrix}\begin{pmatrix}u'\\ v'\end{pmatrix} = \kappa\begin{pmatrix}E & F\\ F & G\end{pmatrix}\begin{pmatrix}u'\\ v'\end{pmatrix},$$

which yields

$$\begin{pmatrix}E & F\\ F & G\end{pmatrix}^{-1}\begin{pmatrix}L & M\\ M & N\end{pmatrix}\begin{pmatrix}u'\\ v'\end{pmatrix} = \kappa\begin{pmatrix}u'\\ v'\end{pmatrix},$$

that is,

$$\frac{1}{EG-F^2}\begin{pmatrix}G & -F\\ -F & E\end{pmatrix}\begin{pmatrix}L & M\\ M & N\end{pmatrix}\begin{pmatrix}u'\\ v'\end{pmatrix} = \kappa\begin{pmatrix}u'\\ v'\end{pmatrix},$$

which yields precisely

$$\begin{aligned}\frac{LG-MF}{EG-F^2}u' + \frac{MG-NF}{EG-F^2}v' &= \kappa u',\\ \frac{ME-LF}{EG-F^2}u' + \frac{NE-MF}{EG-F^2}v' &= \kappa v',\end{aligned}$$

or equivalently

$$\begin{aligned}
\frac{MF-LG}{EG-F^2}u' + \frac{NF-MG}{EG-F^2}v' &= -\kappa u',\\
\frac{LF-ME}{EG-F^2}u' + \frac{MF-NE}{EG-F^2}v' &= -\kappa v'.
\end{aligned}$$

Therefore, the eigenvectors of $J(d\mathbf{N}_{(u,v)})$ correspond to the principal directions at p. $\square$

The fact that $\mathbf{N}_u = -\kappa X_u$ when κ is one of the principal curvatures and when X_u corresponds to the corresponding principal direction (and similarly $\mathbf{N}_v = -\kappa X_v$ for the other principal curvature) is known as the formula of Olinde Rodrigues (1815).

The somewhat irritating negative signs arising in the eigenvalues $-\kappa_1$ and $-\kappa_2$ of $d\mathbf{N}_{(u,v)}$ can be eliminated if we consider the linear map $\mathscr{S}_{(u,v)} = -d\mathbf{N}_{(u,v)}$ instead of $d\mathbf{N}_{(u,v)}$. The map $\mathscr{S}_{(u,v)}$ is called the *shape operator at p*, and the map $d\mathbf{N}_{(u,v)}$ is sometimes called the *Weingarten operator*. The following lemma shows that the second fundamental form arises from the shape operator, and that the shape operator is self-adjoint with respect to the inner product $\langle -, - \rangle$ associated with the first fundamental form.

Lemma 20.5. *Given a surface X, for any point $p = X(u,v)$ on X, the second fundamental form of X at p is given by the formula*

$$II_{(u,v)}(\mathbf{t}) = \langle \mathscr{S}_{(u,v)}(\mathbf{t}), \mathbf{t}\rangle,$$

for every $\mathbf{t} \in \mathbb{R}^2$. The map $\mathscr{S}_{(u,v)} = -d\mathbf{N}_{(u,v)}$ is self-adjoint, that is,

$$\langle \mathscr{S}_{(u,v)}(x), y\rangle = \langle x, \mathscr{S}_{(u,v)}(y)\rangle,$$

for all $x, y \in \mathbb{R}^2$.

Proof. For any tangent vector $\mathbf{t} = X_u x + Y_v y$, since

$$\mathscr{S}_{(u,v)}(X_u x + X_v y) = -d\mathbf{N}_{(u,v)}(X_u x + X_v y) = -\mathbf{N}_u x - \mathbf{N}_v y,$$

we have

$$\begin{aligned}
\langle \mathscr{S}_{(u,v)}(X_u x + X_v y), (X_u x + X_v y)\rangle &= \langle (-\mathbf{N}_u x - \mathbf{N}_v y), (X_u x + X_v y)\rangle\\
&= -(\mathbf{N}_u \cdot X_u)x^2 - (\mathbf{N}_u \cdot X_v + \mathbf{N}_v \cdot X_u)xy\\
&\quad - (\mathbf{N}_v \cdot X_v)y^2.
\end{aligned}$$

However, we already showed in the proof of Lemma 20.3 that

$$\begin{aligned}
L &= \mathbf{N} \cdot X_{uu} = -\mathbf{N}_u \cdot X_u,\\
M &= \mathbf{N} \cdot X_{uv} = -\mathbf{N}_v \cdot X_u,\\
M &= \mathbf{N} \cdot X_{vu} = -\mathbf{N}_u \cdot X_v\\
N &= \mathbf{N} \cdot X_{vv} = -\mathbf{N}_v \cdot X_v,
\end{aligned}$$

and thus that

$$\langle \mathscr{S}_{(u,v)}(X_u x + X_v y),\, (X_u x + X_v y)\rangle = Lx^2 + 2Mxy + Ny^2,$$

the second fundamental form. To prove that $\mathscr{S}_{(u,v)}$ is self-adjoint, it is sufficient to prove it for the basis (X_u, X_v). This amounts to proving that

$$\langle \mathbf{N}_u,\, X_v\rangle = \langle X_u,\, \mathbf{N}_v\rangle.$$

However, we just proved that $\mathbf{N}_v \cdot X_u = \mathbf{N}_u \cdot X_v = -M$, and the proof is complete. $\square$

Thus, in some sense, the shape operator contains all the information about curvature.

Remark: The fact that the first fundamental form I is positive definite and that $\mathscr{S}_{(u,v)}$ is self-adjoint with respect to I can be used to give a fancier proof of the fact that $\mathscr{S}_{(u,v)}$ has two real eigenvalues, that the eigenvectors are orthonormal, and that the eigenvalues correspond to the maximum and the minimum of I on the unit circle. For such a proof, see do Carmo [12]. Our proof is more basic and from first principles.

20.8 The Dupin Indicatrix

The second fundamental form shows up again when we study the deviation of a surface from its tangent plane in a neighborhood of the point of tangency. A way to study this deviation is to imagine that we dip the surface in water, and watch the shorelines formed in the water by the surface in a small region around a chosen point, as we move the surface up and down very gently. The resulting curve is known as the Dupin indicatrix (1813). Formally, consider the tangent plane $T_{(u_0,v_0)}(X)$ at some point $p = X(u_0, v_0)$, and consider the perpendicular distance $\rho(u,v)$ from the tangent plane to a point on the surface defined by (u,v). This perpendicular distance can be expressed as

$$\rho(u,v) = (X(u,v) - X(u_0,v_0)) \cdot \mathbf{N}_{(u_0,v_0)}.$$

However, since X is at least C^3-continuous, by Taylor's formula, in a neighborhood of (u_0, v_0) we can write

$$\begin{aligned} X(u,v) = {} & X(u_0,v_0) + X_u(u-u_0) + X_v(v-v_0) \\ & + \frac{1}{2}\,(X_{uu}(u-u_0)^2 + 2X_{uv}(u-u_0)(v-v_0) + X_{vv}(v-v_0)^2) \\ & + ((u-u_0)^2 + (v-v_0)^2)h_1(u,v), \end{aligned}$$

where $\lim_{(u,v)\to(u_0,v_0)} h_1(u,v) = 0$. However, recall that X_u and X_v are really evaluated at (u_0, v_0) (and so are X_{uu}, $X_{u,v}$, and X_{vv}), and so they are orthogonal to $\mathbf{N}_{(u_0,v_0)}$. From this, dotting with $\mathbf{N}_{(u_0,v_0)}$, we get

$$\begin{aligned}\rho(u,v) = {} & \frac{1}{2}\,(L(u-u_0)^2 + 2M(u-u_0)(v-v_0) + N(v-v_0)^2) \\ & + ((u-u_0)^2 + (v-v_0)^2)h(u,v),\end{aligned}$$

where $\lim_{(u,v)\to(u_0,v_0)} h(u,v) = 0$. Therefore, we get another interpretation of the second fundamental form as a way of measuring the deviation from the tangent plane.

For ε small enough, and in a neighborhood of (u_0, v_0) small enough, the set of points $X(u,v)$ on the surface such that $\rho(u,v) = \pm\frac{1}{2}\varepsilon^2$ will look like portions of the curves of equation

$$\frac{1}{2}\,\left(L(u-u_0)^2 + 2M(u-u_0)(v-v_0) + N(v-v_0)^2\right) = \pm\frac{1}{2}\varepsilon^2.$$

Letting $u - u_0 = \varepsilon x$ and $v - v_0 = \varepsilon y$, these curves are defined by the equations

$$Lx^2 + 2Mxy + Ny^2 = \pm 1.$$

These curves are called the *Dupin indicatrix*. It is more convenient to switch to an orthonormal basis where e_1 and e_2 are eigenvectors of the Gauss map $d\mathbf{N}_{(u_0,v_0)}$. If so, it is immediately seen that

$$Lx^2 + 2Mxy + Ny^2 = \kappa_1 x^2 + \kappa_2 y^2,$$

where κ_1 and κ_2 are the principal curvatures. Thus, the equation of the Dupin indicatrix is

$$\kappa_1 x^2 + \kappa_2 y^2 = \pm 1.$$

There are several cases, depending on the sign of $\kappa_1\kappa_2 = K$, i.e., depending on the sign of $LN - M^2$.

(1) If $LN - M^2 > 0$, then κ_1 and κ_2 have the same sign. This is the case of an *elliptic point*. If $\kappa_1 \neq \kappa_2$, and $\kappa_1 > 0$ and $\kappa_2 > 0$, we get the ellipse of equation

$$\frac{x^2}{\left(\sqrt{\frac{1}{\kappa_1}}\right)^2} + \frac{y^2}{\left(\sqrt{\frac{1}{\kappa_2}}\right)^2} = 1,$$

and if $\kappa_1 < 0$ and $\kappa_2 < 0$, we get the ellipse of equation

$$\frac{x^2}{\left(\sqrt{-\frac{1}{\kappa_1}}\right)^2} + \frac{y^2}{\left(\sqrt{-\frac{1}{\kappa_2}}\right)^2} = 1.$$

When $\kappa_1 = \kappa_2$, i.e., an *umbilical point*, the Dupin indicatrix is a circle.

(2) If $LN - M^2 = 0$ and $L^2 + M^2 + N^2 > 0$, then $\kappa_1 = 0$ or $\kappa_2 = 0$, but not both. This is the case of a *parabolic point*. In this case, the Dupin indicatrix degenerates to two parallel lines, since the equation is either

$$\kappa_1 x^2 = \pm 1$$

or

$$\kappa_2 y^2 = \pm 1.$$

(3) If $LN - M^2 < 0$, then κ_1 and κ_2 have different signs. This is the case of a *hyperbolic point*. In this case, the Dupin indicatrix consists of the two hyperbolae of equations

$$\frac{x^2}{\left(\sqrt{\frac{1}{\kappa_1}}\right)^2} - \frac{y^2}{\left(\sqrt{-\frac{1}{\kappa_2}}\right)^2} = 1$$

if $\kappa_1 > 0$ and $\kappa_2 < 0$, or of equation

$$-\frac{x^2}{\left(\sqrt{-\frac{1}{\kappa_1}}\right)^2} + \frac{y^2}{\left(\sqrt{\frac{1}{\kappa_2}}\right)^2} = 1$$

if $\kappa_1 < 0$ and $\kappa_2 > 0$. These hyperbolae share the same asymptotes, which are the asymptotic directions as defined in Section 20.6, and are given by the equation

$$Lx^2 + 2Mxy + Ny^2 = 0.$$

(4) If $L = M = N$, we have a *planar point*, and in this case, the Dupin indicatrix is undefined.

One should be warned that the Dupin indicatrix for the planar point on the monkey saddle shown in Hilbert and Cohn-Vossen [25], Chapter IV, page 192, is wrong!

Therefore, analyzing the shape of the Dupin indicatrix leads us to rediscover the classification of points on a surface in terms of the principal curvatures. It also lends some intuition to the meaning of the words elliptic, hyperbolic, and parabolic (the last one being a bit misleading). The analysis of $\rho(u,v)$ also shows that in the elliptic case, in a small neighborhood of $X(u,v)$, all points of X are on the same side of the tangent plane. This is like being on the top of a round hill. In the hyperbolic case, in a small neighborhood of $X(u,v)$ there are points of X on both sides of the tangent plane. This is a saddle point or a valley (or mountain pass).

20.9 The *Theorema Egregium* of Gauss, the Equations of Codazzi–Mainardi, and Bonnet's Theorem

In Section 20.5 we expressed the geodesic curvature in terms of the Christoffel symbols, and we also showed that these symbols depend only on E,F,G, i.e., on the first fundamental form. In Section 20.7, we expressed $\mathbf{N}_u$ and $\mathbf{N}_v$ in terms of the coefficients of the first and the second fundamental forms. At first glance, given any six functions E, F, G, L, M, N that are at least C^3-continuous on some open subset U of $\mathbb{R}^2$, and where $E,F > 0$ and $EG - F^2 > 0$, it is plausible that there is a surface X defined on some open subset Ω of U and having $Ex^2 + 2Fxy + Gy^2$ as its first fundamental form and $Lx^2 + 2Mxy + Ny^2$ as its second fundamental form. However, this is false. The problem is that for a surface X, the functions E,F,G,L,M,N are not independent.

In this section we investigate the relations that exist among these functions. We will see that there are three compatibility equations. The first one gives the Gaussian curvature in terms of the first fundamental form only. This is the famous *Theorema Egregium* of Gauss (1827). The other two equations express $M_u - L_v$ and $N_u - M_v$ in terms of L,M,N and the Christoffel symbols. These equations are due to Codazzi (1867) and Mainardi (1856). They were discovered independently by Peterson in 1852 (see Gamkrelidze [20]). Remarkably, these compatibility equations are just what it takes to ensure the existence of a surface (at least locally) with $Ex^2 + 2Fxy + Gy^2$ as its first fundamental form and $Lx^2 + 2Mxy + Ny^2$ as its second fundamental form, an important theorem shown by Ossian Bonnet (1867).

Recall that

$$\begin{aligned} X'' &= X_u u_1'' + X_v u_2'' + X_{uu}(u_1')^2 + 2X_{uv}u_1'u_2' + X_{vv}(u_2')^2, \\ &= (L(u_1')^2 + 2Mu_1'u_2' + N(u_2')^2)\mathbf{N} + \kappa_g \mathbf{n_g}, \end{aligned}$$

and since

$$\kappa_g \mathbf{n_g} = \Bigg(u_1'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma_{i\,j}^{1}\, u_i'u_j'\Bigg) X_u + \Bigg(u_2'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma_{i\,j}^{2}\, u_i'u_j'\Bigg) X_v,$$

we get the equations (due to Gauss)

$$\begin{aligned} X_{uu} &= \Gamma_{1\,1}^{1} X_u + \Gamma_{1\,1}^{2} X_v + L\mathbf{N}, \\ X_{uv} &= \Gamma_{1\,2}^{1} X_u + \Gamma_{1\,2}^{2} X_v + M\mathbf{N}, \\ X_{vu} &= \Gamma_{2\,1}^{1} X_u + \Gamma_{2\,1}^{2} X_v + M\mathbf{N}, \\ X_{vv} &= \Gamma_{2\,2}^{1} X_u + \Gamma_{2\,2}^{2} X_v + N\mathbf{N}, \end{aligned}$$

where the Christoffel symbols $\Gamma_{i\,j}^{k}$ are defined such that

$$\begin{pmatrix} \Gamma_{ij}^1 \\ \Gamma_{ij}^2 \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} [i\,j;1] \\ [i\,j;2] \end{pmatrix},$$

and where

$$\begin{aligned}
[1\,1;1] &= \frac{1}{2}E_u, \quad [1\,1;2] = F_u - \frac{1}{2}E_v, \\
[1\,2;1] &= \frac{1}{2}E_v, \quad [1\,2;2] = \frac{1}{2}G_u, \\
[2\,1;1] &= \frac{1}{2}E_v, \quad [2\,1;2] = \frac{1}{2}G_u, \\
[2\,2;1] &= F_v - \frac{1}{2}G_u, \quad [2\,2;2] = \frac{1}{2}G_v.
\end{aligned}$$

Also, recall from Section 20.7 that we have the Weingarten equations

$$\begin{pmatrix} \mathbf{N}_u \\ \mathbf{N}_v \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} X_u \\ X_v \end{pmatrix} = -\begin{pmatrix} L & M \\ M & N \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} X_u \\ X_v \end{pmatrix}.$$

From the Gauss equations and the Weingarten equations

$$\begin{aligned}
X_{uu} &= \Gamma_{11}^1 X_u + \Gamma_{11}^2 X_v + L\mathbf{N}, \\
X_{uv} &= \Gamma_{12}^1 X_u + \Gamma_{12}^2 X_v + M\mathbf{N}, \\
X_{vu} &= \Gamma_{21}^1 X_u + \Gamma_{21}^2 X_v + M\mathbf{N}, \\
X_{vv} &= \Gamma_{22}^1 X_u + \Gamma_{22}^2 X_v + N\mathbf{N}, \\
\mathbf{N}_u &= aX_u + cX_v, \\
\mathbf{N}_v &= bX_u + dX_v,
\end{aligned}$$

we see that the partial derivatives of X_u, X_v and $\mathbf{N}$ can be expressed in terms of the coefficients E, F, G, L, M, N and their partial derivatives. Thus, a way to obtain relations among these coefficients is to write the equations expressing the commutation of partials, i.e.,

$$\begin{aligned}
(X_{uu})_v - (X_{uv})_u &= 0, \\
(X_{vv})_u - (X_{vu})_v &= 0, \\
\mathbf{N}_{uv} - \mathbf{N}_{vu} &= 0.
\end{aligned}$$

Using the Gauss equations and the Weingarten equations, we obtain relations of the form

$$\begin{aligned}
A_1 X_u + B_1 X_v + C_1 \mathbf{N} &= 0, \\
A_2 X_u + B_2 X_v + C_2 \mathbf{N} &= 0, \\
A_3 X_u + B_3 X_v + C_3 \mathbf{N} &= 0,
\end{aligned}$$

where A_i, B_i, and C_i are functions of E, F, G, L, M, N and their partial derivatives, for $i = 1, 2, 3$. However, since the vectors X_u, X_v, and $\mathbf{N}$ are linearly independent, we obtain the nine equations

$$A_i = 0, \quad B_i = 0, \quad C_i = 0, \quad \text{for } i = 1, 2, 3.$$

Although this is very tedious, it can be shown that these equations are equivalent to just three equations. Due to its importance, we state the *Theorema Egregium* of Gauss.

Theorem 20.1. *Given a surface X and a point $p = X(u, v)$ on X, the Gaussian curvature K at (u, v) can be expressed as a function of E, F, G, and their partial derivatives. In fact,*

$$(EG - F^2)^2 K = \begin{vmatrix} C & F_v - \frac{1}{2} G_u & \frac{1}{2} G_v \\ \frac{1}{2} E_u & E & F \\ F_u - \frac{1}{2} E_v & F & G \end{vmatrix} - \begin{vmatrix} 0 & \frac{1}{2} E_v & \frac{1}{2} G_u \\ \frac{1}{2} E_v & E & F \\ \frac{1}{2} G_u & F & G \end{vmatrix},$$

where

$$C = \frac{1}{2}(-E_{vv} + 2F_{uv} - G_{uu}).$$

Proof. Following Darboux [7] (Volume III, page 246), a way of proving Theorem 20.1 is to start from the formula

$$K = \frac{LN - M^2}{EG - F^2}$$

and to go back to the expressions of L, M, N using D, D', D'' as determinants:

$$L = \frac{D}{\sqrt{EG - F^2}}, \quad M = \frac{D'}{\sqrt{EG - F^2}}, \quad N = \frac{D''}{\sqrt{EG - F^2}},$$

where

$$D = (X_u, X_v, X_{uu}), \quad D' = (X_u, X_v, X_{uv}), \quad D'' = (X_u, X_v, X_{vv}).$$

Then we can write

$$(EG - F^2)^2 K = (X_u, X_v, X_{uu})(X_u, X_v, X_{vv}) - (X_u, X_v, X_{uv})^2,$$

and compute these determinants by multiplying them out. One will eventually get the expression given in the theorem! $\square$

It can be shown that the other two equations, known as the *Codazzi–Mainardi equations*, are the equations

$$M_u - L_v = \Gamma_{11}^2 N - (\Gamma_{12}^2 - \Gamma_{11}^1) M - \Gamma_{12}^1 L,$$
$$N_u - M_v = \Gamma_{12}^2 N - (\Gamma_{22}^2 - \Gamma_{12}^1) M - \Gamma_{22}^1 L.$$

We conclude this section with an important theorem of Ossian Bonnet. First, we show that the first and the second fundamental forms determine a surface up to rigid motion. More precisely, we have the following lemma.

Lemma 20.6. *Let $X\colon \Omega \to \mathbb{E}^3$ and $Y\colon \Omega \to \mathbb{E}^3$ be two surfaces over a connected open set Ω. If X and Y have the same coefficients $E,F,G,\ L,M,N$ over Ω, then there is a rigid motion mapping $X(\Omega)$ onto $Y(\Omega)$.*

The above lemma can be shown using a standard theorem about ordinary differential equations (see do Carmo, [12] Appendix to Chapter 4, pp. 309–314). Finally, we state Bonnet's theorem.

Theorem 20.2. *Let $E,F,G,\ L,M,N$ be C^3-continuous functions on some open set $U \subset \mathbb{R}^2$, and such that $E > 0$, $G > 0$, and $EG - F^2 > 0$. If these functions satisfy the Gauss formula (of the Theorema Egregium) and the Codazzi–Mainardi equations, then for every $(u,v) \in U$ there is an open set $\Omega \subseteq U$ such that $(u,v) \in \Omega$, and a surface $X\colon \Omega \to \mathbb{E}^3$ such that X is a diffeomorphism, and E,F,G are the coefficients of the first fundamental form of X, and L,M,N are the coefficients of the second fundamental form of X. Furthermore, if Ω is connected, then $X(\Omega)$ is unique up to a rigid motion.*

20.10 Lines of Curvature, Geodesic Torsion, Asymptotic Lines

Given a surface X, certain curves on the surface play a special role, for example, the curves corresponding to the directions in which the curvature is maximum or minimum. More precisely, we have the following definition.

Definition 20.8. Given a surface X, a *line of curvature* is a curve $C\colon t \mapsto X(u(t), v(t))$ on X defined on some open interval I and having the property that for every $t \in I$, the tangent vector $C'(t)$ is collinear with one of the principal directions at $X(u(t),v(t))$.

Note that we are assuming that no point on a line of curvature is either a planar point or an umbilical point, since principal directions are undefined as such points. The differential equation defining lines of curvature can be found as follows. Remember from Lemma 20.4 of Section 20.7 that the principal directions are the eigenvectors of $d\mathbf{N}_{(u,v)}$. Therefore, we can find the differential equation defining the lines of curvature by eliminating κ from the two equations from the proof of Lemma 20.4:

$$\frac{MF-LG}{EG-F^2}u' + \frac{NF-MG}{EG-F^2}v' = -\kappa u',$$
$$\frac{LF-ME}{EG-F^2}u' + \frac{MF-NE}{EG-F^2}v' = -\kappa v'.$$

It is not hard to show that the resulting equation can be written as

$$\begin{vmatrix} (v')^2 & -u'v' & (u')^2 \\ E & F & G \\ L & M & N \end{vmatrix} = 0.$$

From the above equation we see that the u-lines and the v-lines are the lines of curvature iff $F = M = 0$. Generally, this differential equation does not have closed-form solutions.

There is another notion that is useful in understanding lines of curvature, the geodesic torsion. Let $C\colon s \mapsto X(u(s), v(s))$ be a curve on X assumed to be parametrized by arc length, and let $X(u(0), v(0))$ be a point on the surface X, and assume that this point is neither a planar point nor an umbilic, so that the principal directions are defined. We can define the orthonormal frame $(e_1, e_2, \mathbf{N})$, known as the *Darboux frame*, where e_1 and e_2 are unit vectors corresponding to the principal directions, $\mathbf{N}$ is the normal to the surface at $X(u(0), v(0))$, and $\mathbf{N} = e_1 \times e_2$.

It is interesting to study the quantity $d\mathbf{N}_{(u,v)}(0)/ds$. If $\mathbf{t} = C'(0)$ is the unit tangent vector at $X(u(0), v(0))$, we have another orthonormal frame considered in Section 20.4, namely $(\mathbf{t}, \mathbf{n_g}, \mathbf{N})$, where $\mathbf{n_g} = \mathbf{N} \times \mathbf{t}$, and if φ is the angle between e_1 and $\mathbf{t}$, we have

$$\begin{aligned} \mathbf{t} &= \cos\varphi\, e_1 + \sin\varphi\, e_2, \\ \mathbf{n_g} &= -\sin\varphi\, e_1 + \cos\varphi\, e_2. \end{aligned}$$

In the following lemma we show that

$$\frac{d\mathbf{N}_{(u,v)}}{ds}(0) = -\kappa_N \mathbf{t} + \tau_g \mathbf{n_g},$$

where κ_N is the normal curvature and where τ_g is a quantity called the *geodesic torsion*.

Lemma 20.7. *Given a curve $C\colon s \mapsto X(u(s), v(s))$ parametrized by arc length on a surface X, we have*

$$\frac{d\mathbf{N}_{(u,v)}}{ds}(0) = -\kappa_N \mathbf{t} + \tau_g \mathbf{n_g},$$

where κ_N is the normal curvature, and where the geodesic torsion τ_g is given by

$$\tau_g = (\kappa_1 - \kappa_2) \sin\varphi \cos\varphi.$$

Proof. Since $-\kappa_1$ and $-\kappa_2$ are the eigenvalues of $d\mathbf{N}_{(u(0),v(0))}$ associated with the eigenvectors e_1 and e_2 (where κ_1 and κ_2 are the principal curvatures), it is immediate that

$$\frac{d\mathbf{N}_{(u,v)}}{ds}(0) = d\mathbf{N}_{(u(0),v(0))}(\mathbf{t}) = -\kappa_1 \cos\varphi e_1 - \kappa_2 \sin\varphi e_2,$$

which shows that this vector is a linear combination of $\mathbf{t}$ and $\mathbf{n_g}$. By projection onto $\mathbf{n_g}$ we get that the geodesic torsion τ_g given by

$$\begin{aligned}\tau_g &= d\mathbf{N}_{(u(0),v(0))}(\mathbf{t})\cdot\mathbf{n_g},\\ &= (-\kappa_1\cos\varphi\, e_1 - \kappa_2\sin\varphi\, e_2)\cdot(-\sin\varphi\, e_1 + \cos\varphi\, e_2),\\ &= (\kappa_1-\kappa_2)\sin\varphi\cos\varphi.\end{aligned}$$

Using Euler's formula (see Section 20.6)

$$\kappa_N = \kappa_1\cos^2\varphi + \kappa_2\sin^2\varphi,$$

it is immediately verified that

$$d\mathbf{N}_{(u(0),v(0))}(\mathbf{t})\cdot\mathbf{t} = -\kappa_N,$$

which proves the lemma. $\square$

From the formula

$$\tau_g = (\kappa_1-\kappa_2)\sin\varphi\cos\varphi,$$

since φ is the angle between the tangent vector to the curve C and a principal direction, it is clear that the lines of curvature are characterized by the fact that $\tau_g = 0$. One will also observe that orthogonal curves have opposite geodesic torsions (same absolute value and opposite signs).

If $\mathbf{N}$ is the principal normal, τ is the torsion of C at $X(u(0),v(0))$, and θ is the angle between $\mathbf{N}$ and $\mathbf{n}$, so that $\cos\theta = \mathbf{N}\cdot\mathbf{n}$, we claim that

$$\tau_g = \tau - \frac{d\theta}{ds},$$

which is often known as *Bonnet's formula*.

Lemma 20.8. *Given a curve $C\colon s\mapsto X(u(s),v(s))$ parametrized by arc length on a surface X, the geodesic torsion τ_g is given by*

$$\tau_g = \tau - \frac{d\theta}{ds} = (\kappa_1-\kappa_2)\sin\varphi\cos\varphi,$$

where τ is the torsion of C at $X(u(0),v(0))$, and θ is the angle between $\mathbf{N}$ and the principal normal $\mathbf{n}$ to C at $s=0$.

Proof. We differentiate

$$\cos\theta = \mathbf{N}\cdot\mathbf{n}.$$

This yields

$$-\sin\theta\frac{d\theta}{ds} = \frac{d\mathbf{N}}{ds}\cdot\mathbf{n} + \mathbf{N}\cdot\frac{d\mathbf{n}}{ds},$$

and since by the Frenet–Serret formulae

$$\frac{d\mathbf{n}}{ds} = -\kappa\mathbf{t} - \tau\mathbf{b},$$

and by Lemma 20.7

$$\frac{d\mathbf{N}}{ds} = -\kappa_N \mathbf{t} + \tau_g \mathbf{n_g},$$

we get

$$\begin{aligned}
-\sin\theta \frac{d\theta}{ds} &= (-\kappa_N \mathbf{t} + \tau_g \mathbf{n_g}) \cdot \mathbf{n} + \mathbf{N} \cdot (-\kappa \mathbf{t} - \tau \mathbf{b}) \\
&= \tau_g (\mathbf{n_g} \cdot \mathbf{n}) - \tau(\mathbf{N} \cdot \mathbf{b}) \\
&= \tau_g \sin\theta - \tau \sin\theta,
\end{aligned}$$

since

$$\mathbf{n_g} \cdot \mathbf{n} = \mathbf{N} \cdot \mathbf{b} = \sin\theta.$$

Therefore, when $\theta \neq 0$, we get

$$-\frac{d\theta}{ds} = \tau_g - \tau,$$

and by continuity, when $\theta = 0$,

$$0 = \tau_g - \tau.$$

Therefore, in all cases we obtain the formula

$$\tau_g = \tau - \frac{d\theta}{ds},$$

which proves the lemma. $\square$

Note that the geodesic torsion depends only on the tangent of curves C. Also, for a curve for which $\theta = 0$, we have $\tau_g = \tau$. Such a curve is also characterized by the fact that the geodesic curvature κ_g is null. As we will see shortly, such curves are called geodesics, which explains the name geodesic torsion for τ_g.

Lemma 20.8 can be used to give a quick proof of a beautiful theorem of Dupin (1813). Dupin's theorem has to do with families of surfaces forming a triply orthogonal system. Given some open subset U of $\mathbb{E}^3$, three families $\mathscr{F}_1$, $\mathscr{F}_2$, $\mathscr{F}_3$ of surfaces form a *triply orthogonal system* for U if for every point $p \in U$ there is a unique surface from each family $\mathscr{F}_i$ passing through p, where $i = 1,2,3$, and any two of these surfaces intersect orthogonally along their curve of intersection. Then Dupin's theorem is as follows.

Theorem 20.3. *The surfaces of a triply orthogonal system intersect each other along lines of curvature.*

Proof. Here is a sketch of the proof. First, we note that if two surfaces X_1 and X_2 intersect along a curve C, and if they form a constant angle along C, then the geodesic torsion τ_g^1 of C on X_1 is equal to the geodesic torsion τ_g^2 of C on X_2. Indeed, if θ_1 is the angle between $\mathbf{N_1}$ and $\mathbf{n}$, and θ_2 is the angle between $\mathbf{N_2}$ and $\mathbf{n}$, where $\mathbf{N_1}$ is the normal to X_1, $\mathbf{N_2}$ is the normal to X_2, and $\mathbf{n}$ is the principal normal to C, then

$$\theta_1 - \theta_2 = \lambda,$$

where λ is some constant, and thus

$$\frac{d\theta_1}{ds} = \frac{d\theta_2}{ds},$$

which shows that

$$\tau_g^1 = \tau - \frac{d\theta_1}{ds} = \tau - \frac{d\theta_2}{ds} = \tau_g^2.$$

Now, if the system of surfaces is triply orthogonal, letting τ_{ij} be the geodesic curvature of the curve of intersection C_{ij} between $X_i \in \mathscr{F}_i$ and $X_j \in \mathscr{F}_j$ (where $1 \leq i < j \leq 3$), which is well defined, since X_i and X_j intersect orthogonally, from a previous observation the geodesic torsions of orthogonal curves are opposite, and thus

$$\tau_{12} = -\tau_{13}, \quad \tau_{23} = -\tau_{12}, \quad \tau_{13} = -\tau_{23},$$

from which we get that

$$\tau_{12} = \tau_{23} = \tau_{13} = 0.$$

However, this means that the curves of intersection are lines of curvature. $\square$

A nice application of Theorem 20.3 is that it is possible to find the lines of curvature on an ellipsoid. Indeed, a system of confocal quadrics is triply orthogonal! (see Berger and Gostiaux [4], Chapter 10, Sections 10.2.2.3, 10.4.9.5, and 10.6.8.3, and Hilbert and Cohn-Vossen [25], Chapter 4, Section 28).

We now turn briefly to asymptotic lines. Recall that asymptotic directions are defined only at points where $K < 0$, and at such points they correspond to the directions for which the normal curvature κ_N is null.

Definition 20.9. Given a surface X, an *asymptotic line* is a curve $C\colon t \mapsto X(u(t), 4v(t))$ on X defined on some open interval I where $K < 0$, and having the property that for every $t \in I$, the tangent vector $C'(t)$ is collinear with one of the asymptotic directions at $X(u(t), v(t))$.

The differential equation defining asymptotic lines is easily found, since it expresses the fact that the normal curvature is null:

$$L(u')^2 + 2M(u'v') + N(v')^2 = 0.$$

Such an equation generally does not have closed-form solutions. Note that the u-lines and the v-lines are asymptotic lines iff $L = N = 0$ (and $F \neq 0$).

Example 20.9. Perseverant readers are welcome to compute E, F, G, L, M, N for the *Enneper surface*

$$x = u - \frac{u^3}{3} + uv^2,$$
$$y = v - \frac{v^3}{3} + u^2 v,$$
$$z = u^2 - v^2.$$

Then they will be able to find closed-form solutions for the lines of curvature and the asymptotic lines.

Parabolic lines are defined by the equation

$$LN - M^2 = 0,$$

where $L^2 + M^2 + N^2 > 0$. In general, the locus of parabolic points consists of several curves and points. For fun, the reader should look at Klein's experiment as described in Hilbert and Cohn-Vossen [25], Chapter IV, Section 29, page 197. We now turn briefly to geodesics.

20.11 Geodesic Lines, Local Gauss–Bonnet Theorem

Geodesics play a very important role in surface theory and in dynamics. One of the main reasons why geodesics are so important is that they generalize to curved surfaces the notion of "shortest path" between two points in the plane (**warning**: As we shall see, this is true only *locally, not globally*). More precisely, given a surface X and any two points $p = X(u_0, v_0)$ and $q = X(u_1, v_1)$ on X, let us look at all the regular curves C on X defined on some open interval I such that $p = C(t_0)$ and $q = C(t_1)$ for some $t_0, t_1 \in I$. It can be shown that in order for such a curve C to minimize the length $l_C(pq)$ of the curve segment from p to q, we must have $\kappa_g(t) = 0$ along $[t_0, t_1]$, where $\kappa_g(t)$ is the geodesic curvature at $X(u(t), v(t))$. In other words, the principal normal $\mathbf{n}$ must be parallel to the normal $\mathbf{N}$ to the surface along the curve segment from p to q. If C is parametrized by arc length, this means that the acceleration must be normal to the surface.

It it then natural to define geodesics as those curves such that $\kappa_g = 0$ everywhere on their domain of definition. Actually, there is another way of defining geodesics in terms of vector fields and covariant derivatives (see do Carmo [12] or Berger and Gostiaux [4]), but for simplicity, we stick to the definition in terms of the geodesic curvature (however, see Section 20.12).

Definition 20.10. Given a surface $X\colon \Omega \to \mathbb{E}^3$, a *geodesic line, or geodesic,* is a regular curve $C\colon I \to \mathbb{E}^3$ on X such that $\kappa_g(t) = 0$ for all $t \in I$.

Note that by regular curve we mean that $\dot{C}(t) \neq 0$ for all $t \in I$, i.e., C is really a curve, and not a single point. Physically, a particle constrained to stay on the surface and not acted on by any force, once set in motion with some nonnull initial velocity (tangent to the surface), will follow a geodesic (assuming no friction).

Since $\kappa_g = 0$ iff the principal normal $\mathbf{n}$ to C at t is parallel to the normal $\mathbf{N}$ to the surface at $X(u(t), v(t))$, and since the principal normal $\mathbf{n}$ is a linear combination of the tangent vector $\dot{C}(t)$ and the acceleration vector $\ddot{C}(t)$, the normal $\mathbf{N}$ to the surface at t belongs to the osculating plane.

The differential equations for geodesics are obtained from Lemma 20.2. Since the tangential part of the curvature at a point is given by

$$\kappa_g \mathbf{n_g} = \Bigg(u_1'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma_{i\,j}^{1}\, u_i' u_j'\Bigg) X_u + \Bigg(u_2'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma_{i\,j}^{2}\, u_i' u_j'\Bigg) X_v,$$

the differential equations for geodesics are

$$\begin{aligned} u_1'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma_{i\,j}^{1}\, u_i' u_j' &= 0, \\ u_2'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma_{i\,j}^{2}\, u_i' u_j' &= 0, \end{aligned}$$

or more explicitly (letting $u = u_1$ and $v = u_2$),

$$\begin{aligned} u'' + \Gamma_{1\,1}^{1}\,(u')^2 + 2\Gamma_{1\,2}^{1}\,u'v' + \Gamma_{2\,2}^{1}\,(v')^2 &= 0, \\ v'' + \Gamma_{1\,1}^{2}\,(u')^2 + 2\Gamma_{1\,2}^{2}\,u'v' + \Gamma_{2\,2}^{2}\,(v')^2 &= 0. \end{aligned}$$

In general, it is impossible to find closed-form solutions for these equations. Nevertheless, from the theory of ordinary differential equations, the following lemma showing the local existence of geodesics can be shown (see do Carmo [12], Chapter 4, Section 4.7).

Lemma 20.9. *Given a surface X, for every point $p = X(u,v)$ on X and every nonnull tangent vector $v \in T_{(u,v)}(X)$, there is some $\varepsilon > 0$ and a unique curve $\gamma \colon\,]-\varepsilon, \varepsilon[\to \mathbb{E}^3$ on the surface X such that γ is a geodesic, $\gamma(0) = p$, and $\gamma'(0) = v$.*

To emphasize that the geodesic γ depends on the initial direction v, we often write $\gamma(t,v)$ instead of $\gamma(t)$. The geodesics on a sphere are the great circles (the plane sections by planes containing the center of the sphere). More generally, in the case of a surface of revolution (a surface generated by a plane curve rotating around an axis in the plane containing the curve and not meeting the curve), the differential equations for geodesics can be used to study the geodesics.

Example 20.10. For example, the meridians are geodesics (meridians are the plane sections by planes through the axis of rotation: They are obtained by rotating the original curve generating the surface). Also, the parallel circles such that at every point p the tangent to the meridian through p is parallel to the axis of rotation is a geodesic. In general, there are other geodesics. For more on geodesics on surfaces of revolution, see do Carmo [12], Chapter 4, Section 4, and the problems.

The geodesics on an ellipsoid are also fascinating; see Berger and Gostiaux [4], Section 10.4.9.5, and Hilbert and Cohn-Vossen [25], Chapter 4, Section 32.

It should be noted that geodesics can be self-intersecting or closed. A deeper study of geodesics requires a study of vector fields on surfaces and would lead us too far. Technically, what is needed is the exponential map, which we now discuss briefly.

The idea behind the exponential map is to parametrize locally the surface X in terms of a map from the tangent space to the surface, this map being defined in terms of short geodesics. More precisely, for every point $p = X(u,v)$ on the surface, there is some open disk B_ε of center $(0,0)$ in $\mathbb{R}^2$ (recall that the tangent plane $T_p(X)$ at p is isomorphic to $\mathbb{R}^2$) and an injective map

$$\exp_p \colon B_\varepsilon \to X(\Omega)$$

such that for every $v \in B_\varepsilon$ with $v \neq 0$,

$$\exp_p(v) = \gamma(1,v),$$

where $\gamma(t,v)$ is the unique geodesic segment such that $\gamma(0,v) = p$ and $\gamma'(0,v) = v$. Furthermore, for B_ε small enough, $\exp_p$ is a diffeomorphism. It turns out that $\exp_p(v)$ is the point q obtained by "laying off" a length equal to $\|v\|$ along the unique geodesic that passes through p in the direction v. Of course, to make sure that all this is well-defined, it is necessary to prove a number of facts. We state the following lemmas, whose proofs can be found in do Carmo [12].

Lemma 20.10. *Given a surface $X \colon \Omega \to \mathbb{E}^3$, for every $v \neq 0$ in $\mathbb{R}^2$, if*

$$\gamma(-,v) \colon\ ]-\varepsilon, \varepsilon\,[\to \mathbb{E}^3$$

is a geodesic on the surface X, then for every $\lambda > 0$, the curve

$$\gamma(-,\lambda v) \colon\ ]-\varepsilon/\lambda, \varepsilon/\lambda\,[\to \mathbb{E}^3$$

is also a geodesic, and

$$\gamma(t,\lambda v) = \gamma(\lambda t, v).$$

From Lemma 20.10, for $v \neq 0$, if $\gamma(1,v)$ is defined, then

$$\gamma\left(\|v\|, \frac{v}{\|v\|}\right) = \gamma(1,v).$$

This leads to the definition of the exponential map.

Definition 20.11. Given a surface $X \colon \Omega \to \mathbb{E}^3$ and a point $p = X(u,v)$ on X, the *exponential map* $\exp_p$ is the map

$$\exp_p \colon U \to X(\Omega)$$

defined such that

$$\exp_p(v) = \gamma\left(\|v\|, \frac{v}{\|v\|}\right) = \gamma(1, v),$$

where $\gamma(0,v) = p$ and U is the open subset of $\mathbb{R}^2 (= T_p(X))$ such that for every $v \neq 0$, $\gamma(\|v\|, v/\|v\|)$ is defined. We let $\exp_p(0) = p$.

It is immediately seen that U is star-like. One should realize that in general, U is a proper subset of Ω. For example, in the case of a sphere, the exponential map is defined everywhere. However, given a point p on a sphere, if we remove its antipodal point $-p$, then $\exp_p(v)$ is undefined for points on the circle of radius π. Nevertheless, $\exp_p$ is always well-defined in a small open disk.

Lemma 20.11. *Given a surface $X\colon \Omega \to \mathbb{E}^3$, for every point $p = X(u,v)$ on X there is some $\varepsilon > 0$, some open disk B_ε of center $(0,0)$, and some open subset V of $X(\Omega)$ with $p \in V$ such that the exponential map $\exp_p\colon B_\varepsilon \to V$ is well-defined and is a diffeomorphism.*

A neighborhood of p on X of the form $\exp_p(B_\varepsilon)$ is called a *normal neighborhood of p*. The exponential map can be used to define special local coordinate systems on normal neighborhoods, by picking special coordinate systems on the tangent plane. In particular, we can use polar coordinates (ρ, θ) on $\mathbb{R}^2$. In this case, $0 < \theta < 2\pi$. Thus, the closed half-line corresponding to $\theta = 0$ is omitted, and so is its image under $\exp_p$. It is easily seen that in such a coordinate system $E = 1$ and $F = 0$, and ds^2 is of the form

$$ds^2 = dr^2 + G\,d\theta^2.$$

The image under $\exp_p$ of a line through the origin in $\mathbb{R}^2$ is called a *geodesic line*, and the image of a circle centered at the origin is called a *geodesic circle*. Since $F = 0$, these lines are orthogonal. It can also be shown that the Gaussian curvature is expressed as follows:

$$K = -\frac{1}{\sqrt{G}} \frac{\partial^2(\sqrt{G})}{\partial \rho^2}.$$

Polar coordinates can be used to prove the following lemma showing that geodesics locally minimize arc length.

However, globally, geodesics generally do not minimize arc length. For instance, on a sphere, given any two nonantipodal points p, q, since there is a unique great circle passing through p and q, there are two geodesic arcs joining p and q, but only one of them has minimal length.

Lemma 20.12. *Given a surface $X\colon \Omega \to \mathbb{E}^3$, for every point $p = X(u,v)$ on X there is some $\varepsilon > 0$ and some open disk B_ε of center $(0,0)$ such that for every $q \in \exp_p(B_\varepsilon)$ and geodesic $\gamma\colon\ ]-\eta, \eta[\to \mathbb{E}^3$ in $\exp_p(B_\varepsilon)$ such that $\gamma(0) = p$ and $\gamma(t_1) = q$, and for every regular curve $\alpha\colon [0, t_1] \to \mathbb{E}^3$ on X such that $\alpha(0) = p$ and $\alpha(t_1) = q$, we have*

$$l_\gamma(pq) \leq l_\alpha(pq),$$

where $l_\alpha(pq)$ denotes the length of the curve segment α from p to q (and similarly for γ). Furthermore, $l_\gamma(pq) = l_\alpha(pq)$ iff the trace of γ is equal to the trace of α between p and q.

As we already noted, Lemma 20.12 is false globally, since a geodesic, if extended too much, may not be the shortest path between two points (example of the sphere). However, the following lemma shows that a shortest path must be a geodesic segment.

Lemma 20.13. *Given a surface $X\colon \Omega \to \mathbb{E}^3$, let $\alpha\colon I \to \mathbb{E}^3$ be a regular curve on X parametrized by arc length. For any two points $p = \alpha(t_0)$ and $q = \alpha(t_1)$ on α, assume that the length $l_\alpha(pq)$ of the curve segment from p to q is minimal among all regular curves on X passing through p and q. Then α is a geodesic.*

At this point, in order to go further into the theory of surfaces, in particular closed surfaces, it is necessary to introduce differentiable manifolds and more topological tools. However, this is beyond the scope of this book, and we simply refer the interested readers to the following sources. For the foundations of differentiable manifolds, see Berger and Gostiaux [4], do Carmo [12, 13, 14], Guillemin and Pollack [24], Warner [43], Sternberg [41], Boothby [5], Lafontaine [29], Lehmann and Sacré [31], Gray [23], Stoker [42], Gallot, Hulin, and Lafontaine [19], Milnor [36], Lang [30], Malliavin [33], and Godbillon [21]. Abraham and Marsden [1] contains a compact and yet remarkably clear and complete presentation of differentiable manifolds and Riemannian geometry (and a lot of Lagrangian and Hamiltonian mechanics!). For the differential topology of surfaces, see Guillemin and Pollack [24], Milnor [36, 37], Hopf [26], Gramain [22], Lehmann and Sacré [31], and for the algebraic topology of surfaces, see Chapter 1 of Massey [35, 34] and Chapter 1 of Ahlfors and Sario [2], which is remarkable. A lively and remarkably clear introduction to algebraic topology, including the classification theorem for surfaces, can be found in Fulton [17]. For a detailed presentation of differential geometry and Riemannian geometry, see do Carmo [14], Gallot, Hulin, and Lafontaine [19], Sternberg [41], Gray [23], Sharpe [40], Lang [30], Lehmann and Sacré [31], and Malliavin [33]. Choquet-Bruhat [6] also covers a lot of geometric analysis, differential geometry, and topology, and stresses applications to physics. Volume 28 of the *Encyclopaedia of Mathematical Sciences* edited by Gamkrelidze [20] contains a very interesting survey of the field of differential geometry, understood in a broad sense.

Nevertheless, we cannot resist to state one of the "gems" of the differential geometry of surfaces, the local Gauss–Bonnet theorem.

The local Gauss–Bonnet theorem deals with regions on a surface homeomorphic to a closed disk whose boundary is a closed piecewise regular curve α without self-intersection. Such a curve has a finite number of points where the tangent has a discontinuity. If there are n such discontinuities $p_1, \ldots, p_n$, let θ_i be the exterior angle between the two tangents at p_i. More precisely, if $\alpha(t_i) = p_i$, and the two tangents at p_i are defined by the vectors

$$\lim_{t \to t_i, t < t_i} \alpha'(t) = \alpha'_-(t_i) \neq 0,$$

and

$$\lim_{t \to t_i, t > t_i} \alpha'(t) = \alpha'_+(t_i) \neq 0,$$

the angle θ_i is defined as follows. Let θ_i be the angle between $\alpha'_-(t_i)$ and $\alpha'_+(t_i)$ such that $0 < |\theta_i| \leq \pi$, its sign being determined as follows. If p_i is not a cusp, which means that $|\theta_i| \neq \pi$, we give θ_i the sign of the determinant

$$(\alpha'_-(t_i),\, \alpha'_+(t_i),\, \mathbf{N}_{\mathbf{p_i}}).$$

If p_i is a cusp, which means that $|\theta_i| = \pi$, it is easy to see that there is some $\varepsilon > 0$ such that the determinant

$$(\alpha'(t_i - \eta),\, \alpha'(t_i + \eta),\, \mathbf{N}_{\mathbf{p_i}})$$

does not change sign for $\eta \in]-\varepsilon, \varepsilon[$, and we give θ_i the sign of this determinant. Let us call a region defined as above a *simple region*. In order to state a simpler version of the theorem, let us also assume that the curve segments between consecutive points p_i are geodesic lines. We will call such a curve a *geodesic polygon*. Then the *local Gauss–Bonnet theorem* can be stated as follows.

Theorem 20.4. *Given a surface $X\colon \Omega \to \mathbb{E}^3$, assuming that X is injective, $F = 0$, and that Ω is an open disk, for every simple region R of $X(\Omega)$ bounded by a geodesic polygon with n vertices $p_1, \ldots, p_n$, letting $\theta_1, \ldots, \theta_n$ be the exterior angles of the geodesic polygon, we have*

$$\iint_R K dA + \sum_{i=1}^{n} \theta_i = 2\pi.$$

Remark: The assumption that $F = 0$ is not essential, it simply makes the proof easier.

Some clarification regarding the meaning of the integral $\iint_R K dA$ is in order. Firstly, it can be shown that the element of area dA on a surface X is given by

$$dA = \|X_u \times X_v\| du\, dv = \sqrt{EG - F^2}\, du\, dv.$$

Secondly, if we recall from Lemma 20.3 that

$$\begin{pmatrix} \mathbf{N}_u \\ \mathbf{N}_v \end{pmatrix} = -\begin{pmatrix} L & M \\ M & N \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} X_u \\ X_v \end{pmatrix},$$

it is easily verified that

$$\mathbf{N}_u \times \mathbf{N}_v = \frac{LN - M^2}{EG - F^2} X_u \times X_v = K(X_u \times X_v).$$

Thus,

$$\iint_R K dA = \iint_R K\|X_u \times X_v\| du\, dv = \iint_R \|\mathbf{N}_u \times \mathbf{N}_v\| du\, dv,$$

the latter integral representing the area of the spherical image of R under the Gauss map. This is the interpretation of the integral $\iint_R K dA$ that Gauss himself gave.

If the geodesic polygon is a triangle, and if A, B, C are the interior angles, so that $A = \pi - \theta_1$, $B = \pi - \theta_2$, $C = \pi - \theta_3$, the Gauss–Bonnet theorem reduces to what is known as the *Gauss formula*:

$$\iint_R K dA = A + B + C - \pi.$$

The above formula shows that if $K > 0$ on R, then $\iint_R K dA$ is the excess of the sum of the angles of the geodesic triangle over π. If $K < 0$ on R, then $\iint_R K dA$ is the defficiency of the sum of the angles of the geodesic triangle over π. And finally, if $K = 0$, then $A + B + C = \pi$, which we know from the plane!

For the global version of the Gauss–Bonnet theorem, we need the topological notion of the Euler–Poincaré characteristic. If S is an orientable compact surface with g holes, the *Euler–Poincaré characteristic* $\chi(S)$ of S is defined by

$$\chi(S) = 2(1 - g).$$

Then the Gauss–Bonnet theorem states that

$$\iint_S K dA = 2\pi\chi(S).$$

What is remarkable about the above formula is that it relates the topology of the surface (its *genus* g, the number of holes) and the geometry of S, i.e., how it curves. However, all this is beyond the scope of this book. For more information the interested reader is referred to Berger and Gostiaux [4], do Carmo [12, 13, 14], Hopf [26], Milnor [36], Lehmann and Sacré [31], Chapter 1 of Massey [35, 34], Chapter 1 of Ahlfors and Sario [2], and Fulton [17].

20.12 Covariant Derivative, Parallel Transport, Geodesics Revisited

Another way to approach geodesics is in terms of covariant derivatives. The notion of covariant derivative is a key concept of Riemannian geometry, and this section provides a down-to-earth presentation of this notion in the case of a surface.

Let $X\colon \Omega \to \mathbb{E}^3$ be a surface. Given any open subset U of X, a *vector field on* U is a function w that assigns to every point $p \in U$ some tangent vector $w(p) \in T_pX$ to X at p. A vector field w on U is *differentiable at* p if when expressed as $w = aX_u + bX_v$

in the basis (X_u, X_v) (of T_pX), the functions a and b are differentiable at p. A vector field w is *differentiable on* U when it is differentiable at every point $p \in U$.

Definition 20.12. Let w be a differentiable vector field on some open subset U of a surface X. For every $y \in T_pX$, consider a curve $\alpha\colon\ ]-\varepsilon,\varepsilon[\to U$ on X with $\alpha(0) = p$ and $\alpha'(0) = y$, and let $w(t) = (w \circ \alpha)(t)$ be the restriction of the vector field w to the curve α. The normal projection of $dw/dt(0)$ onto the plane T_pX, denoted by

$$\frac{Dw}{dt}(0), \quad \text{or} \quad D_{\alpha'}w(p), \quad \text{or} \quad D_y w(p),$$

is called the *covariant derivative of* w *at* p *relative to* y.

The definition of $Dw/dt(0)$ seems to depend on the curve α, but in fact, it depends only on y and the first fundamental form of X. Indeed, if $\alpha(t) = X(u(t), v(t))$, from

$$w(t) = a(u(t), v(t))X_u + b(u(t), v(t))X_v,$$

we get

$$\frac{dw}{dt} = a(X_{uu}\dot{u} + X_{uv}\dot{v}) + b(X_{vu}\dot{u} + X_{vv}\dot{v}) + \dot{a}X_u + \dot{b}X_v.$$

However, we obtained earlier the following formula (due to Gauss) for X_{uu}, X_{uv}, X_{vu}, and X_{vv}:

$$\begin{aligned}
X_{uu} &= \Gamma_{11}^1 X_u + \Gamma_{11}^2 X_v + L\mathbf{N},\\
X_{uv} &= \Gamma_{12}^1 X_u + \Gamma_{12}^2 X_v + M\mathbf{N},\\
X_{vu} &= \Gamma_{21}^1 X_u + \Gamma_{21}^2 X_v + M\mathbf{N},\\
X_{vv} &= \Gamma_{22}^1 X_u + \Gamma_{22}^2 X_v + N\mathbf{N}.
\end{aligned}$$

Now Dw/dt is the tangential component of dw/dt. Thus by dropping the normal components, we get

$$\begin{aligned}
\frac{Dw}{dt} &= (\dot{a} + \Gamma_{11}^1 a\dot{u} + \Gamma_{12}^1 a\dot{v} + \Gamma_{21}^1 b\dot{u} + \Gamma_{22}^1 b\dot{v})X_u\\
&\quad + (\dot{b} + \Gamma_{11}^2 a\dot{u} + \Gamma_{12}^2 a\dot{v} + \Gamma_{21}^2 b\dot{u} + \Gamma_{22}^2 b\dot{v})X_v.
\end{aligned}$$

Thus, the covariant derivative depends only on $y = (\dot{u}, \dot{v})$ and the Christoffel symbols, but we know that those depend only on the first fundamental form of X.

Definition 20.13. Let $\alpha\colon I \to X$ be a regular curve on a surface X. A *vector field along* α is a map w that assigns to every $t \in I$ a vector $w(t) \in T_{\alpha(t)}X$ in the tangent plane to X at $\alpha(t)$. Such a vector field is differentiable if the components a, b of $w = aX_u + bX_v$ over the basis (X_u, X_v) are differentiable. The expression $Dw/dt(t)$ defined in the above equation is called the *covariant derivative of* w *at* t.

Definition 20.13 extends immediately to piecewise regular curves on a surface.

Definition 20.14. Let $\alpha\colon I \to X$ be a regular curve on a surface X. A vector field along α is *parallel* if $Dw/dt = 0$ for all $t \in I$.

Thus, a vector field along a curve on a surface is parallel iff its derivative is normal to the surface. For example, if C is a great circle on the sphere S^2 parametrized by arc length, the vector field of tangent vectors $C'(s)$ along C is a parallel vector field. We get the following alternative definition of a geodesic.

Definition 20.15. Let $\alpha\colon I \to X$ be a nonconstant regular curve on a surface X. Then α is a *geodesic* if the field of its tangent vectors $\dot{\alpha}(t)$ is parallel along α, that is,

$$\frac{D\dot{\alpha}}{dt}(t) = 0$$

for all $t \in I$.

If we let $\alpha(t) = X(u(t), v(t))$, from the equation

$$\begin{aligned}\frac{Dw}{dt} &= (\dot{a} + \Gamma_{1\,1}^{1} a\dot{u} + \Gamma_{1\,2}^{1} a\dot{v} + \Gamma_{2\,1}^{1} b\dot{u} + \Gamma_{2\,2}^{1} b\dot{v})X_u \\ &\quad + (\dot{b} + \Gamma_{1\,1}^{2} a\dot{u} + \Gamma_{1\,2}^{2} a\dot{v} + \Gamma_{2\,1}^{2} b\dot{u} + \Gamma_{2\,2}^{2} b\dot{v})X_v,\end{aligned}$$

with $a = \dot{u}$ and $b = \dot{v}$, we get the equations

$$\begin{aligned}\ddot{u} + \Gamma_{1\,1}^{1}(\dot{u})^2 + \Gamma_{1\,2}^{1}\dot{u}\dot{v} + \Gamma_{2\,1}^{1}\dot{u}\dot{v} + \Gamma_{2\,2}^{1}(\dot{v})^2 &= 0,\\ \ddot{v} + \Gamma_{1\,1}^{2}(\dot{u})^2 + \Gamma_{1\,2}^{2}\dot{u}\dot{v} + \Gamma_{2\,1}^{2}\dot{u}\dot{v} + \Gamma_{2\,2}^{2}(\dot{v})^2 &= 0,\end{aligned}$$

which are indeed the equations of geodesics found earlier, since $\Gamma_{1\,2}^{1} = \Gamma_{2\,1}^{1}$ and $\Gamma_{1\,2}^{2} = \Gamma_{2\,1}^{2}$ (except that α is not necessarily parametrized by arc length).

Lemma 20.14. *Let $\alpha\colon I \to X$ be a regular curve on a surface X, and let v and w be two parallel vector fields along α. Then the inner product $\langle v(t), w(t)\rangle$ is constant along α (where $\langle -,-\rangle$ is the inner product associated with the first fundamental form, i.e., the Riemannian metric). In particular, $\|v\|$ and $\|w\|$ are constant and the angle between $v(t)$ and $w(t)$ is also constant.*

Proof. The vector field $v(t)$ is parallel iff dv/dt is normal to the tangent plane to the surface X at $\alpha(t)$, and so

$$\langle v'(t), w(t)\rangle = 0$$

for all $t \in I$. Similarly, since $w(t)$ is parallel, we have

$$\langle v(t), w'(t)\rangle = 0$$

for all $t \in I$. Then

$$\langle v(t), w(t)\rangle' = \langle v'(t), w(t)\rangle + \langle v(t), w'(t)\rangle = 0$$

for all $t \in I$, which means that $\langle v(t), w(t)\rangle$ is constant along α. $\square$

As a consequence of Corollary 20.14, if $\alpha\colon I \to X$ is a nonconstant geodesic on X, then $\|\dot{\alpha}\| = c$ for some constant $c > 0$. Thus, we may reparametrize α with respect to the arc length $s = ct$, and we note that the parameter t of a geodesic is proportional to the arc length of α.

Lemma 20.15. *Let $\alpha\colon I \to X$ be a regular curve on a surface X, and for any $t_0 \in I$, let $w_0 \in T_{\alpha(t_0)}X$. Then there is a unique parallel vector field $w(t)$ along α such that $w(t_0) = w_0$.*

Lemma 20.15 is an immediate consequence of standard results on ODEs. This lemma yields the notion of parallel transport.

Definition 20.16. Let $\alpha\colon I \to X$ be a regular curve on a surface X, and for any $t_0 \in I$, let $w_0 \in T_{\alpha(t_0)}X$. Let w be the parallel vector field along α, so that $w(t_0) = w_0$, given by Lemma 20.15. Then for any $t \in I$, the vector $w(t)$ is called the *parallel transport of w_0 along α at t*.

It is easily checked that the parallel transport does not depend on the parametrization of α. If X is an open subset of the plane, then the parallel transport of w_0 at t is indeed a vector $w(t)$ parallel to w_0 (in fact, equal to w_0). However, on a curved surface, the parallel transport may be somewhat counterintuitive.

If two surfaces X and Y are tangent along a curve $\alpha\colon I \to X$, and if $w_0 \in T_{\alpha(t_0)}X = T_{\alpha(t_0)}Y$ is a tangent vector to both X and Y at t_0, then the parallel transport of w_0 along α is the same whether it is relative to X or relative to Y. This is because Dw/dt is the same for both surfaces, and by uniqueness of the parallel transport, the assertion follows. This property can be used to figure out the parallel transport of a vector w_0 when Y is locally isometric to the plane.

In order to generalize the notion of covariant derivative, geodesic, and curvature to manifolds more general than surfaces, the notion of *connection* is needed.

If M is a manifold, we can consider the space $\mathscr{X}(M)$ of smooth vector fields X on M. They are smooth maps that assign to every point $p \in M$ some vector $X(p)$ in the tangent space T_pM to M at p. We can also consider the set $\mathscr{C}^\infty(M)$ of smooth functions $f\colon M \to \mathbb{R}$ on M. Then an *affine connection D on M* is a differentiable map

$$D\colon \mathscr{X}(M) \times \mathscr{X}(M) \to \mathscr{X}(M),$$

denoted by $D_X Y$ (or $\nabla_X Y$), satisfying the following properties:

(1) $D_{fX+gY}Z = fD_XZ + gD_YZ$;
(2) $D_X(\lambda Y + \mu Z) = \lambda D_X Y + \mu D_X Z$;
(3) $D_X(fY) = fD_XY + X(f)Y$,

for all $\lambda, \mu \in \mathbb{R}$, all $X, Y, Z \in \mathscr{X}(M)$, and all $f, g \in \mathscr{C}^\infty(M)$, where $X(f)$ denotes the directional derivative of f in the direction X.

Thus, an affine connection is $\mathscr{C}^\infty(M)$-linear in X, $\mathbb{R}$-linear in Y, and satisfies a "Leibniz"-type law in Y. For any chart $\varphi\colon U \to \mathbb{R}^m$, denoting the coordinate functions by $x_1, \ldots, x_m$, if X is given locally by

$$X(p) = \sum_{i=1}^{m} a_i(p) \frac{\partial}{\partial x_i},$$

then

$$X(f)(p) = \sum_{i=1}^{m} a_i(p) \frac{\partial (f \circ \varphi^{-1})}{\partial x_i}.$$

It can be checked that $X(f)$ does not depend on the choice of chart.

The intuition behind a connection is that $D_X Y$ is the directional derivative of Y in the direction X. The notion of covariant derivative can be introduced via the following lemma:

Lemma 20.16. *Let M be a smooth manifold and assume that D is an affine connection on M. Then there is a unique map D associating with every vector field V along a curve $\alpha \colon I \to M$ on M another vector field DV/dt along c (the covariant derivative of V along c), such that:*

(1)

$$\frac{D}{dt}(\lambda V + \mu W) = \lambda \frac{DV}{dt} + \mu \frac{DW}{dt},$$

(2)

$$\frac{D}{dt}(fV) = \frac{df}{dt} V + f \frac{DV}{dt},$$

(3) if V is induced by a vector field $Y \in \mathscr{X}(M)$, in the sense that $V(t) = Y(\alpha(t))$, then

$$\frac{DV}{dt} = D_{\alpha'(t)} Y.$$

Then in local coordinates, DV/dt can be expressed in terms of the Chistoffel symbols, pretty much as in the case of surfaces. Parallel vector fields, parallel transport, geodesics, are defined as before.

Affine connections are uniquely induced by Riemannian metrics, a fundamental result of Levi-Civita. In fact, such connections are *compatible with the metric*, which means that for any smooth curve α on M and any two parallel vector fields X, Y along α, the inner product $\langle X, Y \rangle$ is constant. Such connections are also *symmetric*, which means that

$$D_X Y - D_Y X = [X, Y],$$

where $[X, Y]$ is the Lie bracket of vector fields.

For more on all this, consult do Carmo [12, 13], Gallot, Hulin, and Lafontaine [19], or any other text on Riemannian geometry.

20.13 Applications

We saw in Section 19.11 that many engineering problems can be reduced to finding curves having some desired properties. Surfaces also play an important role in engineering problems where modeling 3D shapes is required. Again, this is true of computer graphics and medical imaging, where 3D contours of shapes, for instance organs, are modeled as surfaces. As in the case of curves, in most practical applications it is necessary to consider surfaces composed of various patches, and the problem then arises to join these patches as smoothly as possible, without restricting too much the number of degrees of freedom required for the design. Various kinds of *spline surfaces* were invented to solve this problem. But this time, the situation is more complex than in the case of curves, because there are two kinds of surface patches, rectangular and triangular. Roughly speaking, since rectangular patches are basically products of curves, their spline theory is rather well understood. This is not the case for triangular patches, for which the theory of splines is very sparse (Loop [32] being a noteworthy exception). Thus, we will restrict our brief discussion to rectangular patches. As for curves, there is a notion of *parametric* C^n*-continuity* and of B*-spline*. The theory of B-splines is quite extensive. Among the many references, we recommend Farin [16, 15], Hoschek and Lasser [27], Bartels, Beatty, and Barsky [3], Piegl and Tiller [39], or Gallier [18]. However, since parametric continuity is sometimes too constraining, more flexible continuity conditions have been investigated. There are various notions of *geometric continuity*, or G^n*-continuity*. Roughly speaking, two surface patches join with G^n-continuity if there is a reparametrization (a diffeomorphism) after which the patches join with parametric C^n-continuity along the common boundary curve. As a consequence, geometric continuity may be defined using the chain rule, in terms of a certain *connection matrix*.

One of the most important applications of geometric continuity occurs when two or more rectangular patches are stitched together. In such cases polygonal holes can occur between patches. It is often impossible to fill these holes with patches that join with parametric continuity, and a geometrically continuous solution must be used instead. There are also variations on the theme of geometric continuity, which seems to be a topic of current interest. Again, we refer the readers to Farin [16, 15], Hoschek and Lasser [27], Bartels, Beatty, and Barsky [3], Piegl and Tiller [39], and Loop [32].

As in the case of curves, traditional methods for surface design focus on achieving a specific level of interelement continuity, but the resulting shapes often possess bulges and undulations, and thus are of poor quality. They lack *fairness*. Fairness refers to the quality of regularity of the curvature of a surface. The maximum rate of change of curvature should be minimized. This suggests several approaches.

- Minimal energy surface (which bends as little as possible): Minimize

$$\int_S \left(\kappa_1^2 + \kappa_2^2\right) dA$$

  where κ_1 and κ_2 are the principal curvatures.
- Minimal variation surface (which bends as smoothly as possible): Minimize

$$\int_S \left[(\mathrm{D}_{e_1}\kappa_1)^2 + (\mathrm{D}_{e_2}\kappa_2)^2\right] dA$$

where κ_1 and κ_2 are the principal curvatures and e_1 and e_2 are unit vectors giving the principal directions.

As in the case of curves, these problems can be cast as constrained optimization problems. More details on this approach, called *variational surface design*, can be found in the Ph.D. theses of Henry Moreton [38] and William Welch [44].

20.14 Problems

20.1. Consider the surface X defined by

$$x = v\cos u,$$
$$y = v\sin u,$$
$$z = v.$$

(i) Show that F is regular at every point (u,v), except when $v = 0$.
(ii) Show that X is the set of points such that

$$x^2 + y^2 = z^2.$$

What does this surface look like?

20.2. Let $\alpha\colon I \to \mathbb{E}^3$ be a regular curve whose curvature is nonzero for all $t \in I$, where $I =]a,b[$. Let X be the surface defined over $I \times \mathbb{R}$ such that

$$X(u,v) = \alpha(u) + v\alpha'(u).$$

Show that X is regular for all (u,v) where $v \neq 0$.

Remark: The surface X is called the *tangent surface of* α. The curve α is a *line of striction* on X.

20.3. Let $\alpha\colon I \to \mathbb{E}^3$ be a regular curve whose curvature is nonzero for all $t \in I$, where $I =]a,b[$, and assume that α is parametrized by arc length. For any $r > 0$, let X be the surface defined over $I \times \mathbb{R}$ such that

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

where $(\mathbf{t},\mathbf{n},\mathbf{b})$ is the Frenet frame of α at u.

Show that for every (u,v) such that $X(u,v)$ is regular, the unit normal vector $\mathbf{N}_{(u,v)}$ to X at (u,v) is given by

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

Remark: The surface X is called the *tube of radius r around* α.

20.4. (i) Show that the normals to a regular surface defined by

$$\begin{aligned}
x &= f(v)\cos u,\\
y &= f(v)\sin u,\\
z &= g(v),
\end{aligned}$$

all pass through the z-axis.

Remark: Such a surface is called a *surface of revolution*.

(ii) If S is a connected regular surface and all its normals meet the z axis, show that S has a parametrization as in (i).

20.5. Show that the first fundamental form of a plane and the first fundamental form of a cylinder of revolution defined by

$$X(u,v) = (\cos u, \sin u, v)$$

are both $(E,F,G) = (1,0,1)$.

20.6. Given a helicoid defined such that

$$\begin{aligned}
x &= u_1 \cos v_1,\\
y &= u_1 \sin v_1,\\
z &= v_1,
\end{aligned}$$

show that $(E,F,G) = (1,0,u_1^2+1)$.

20.7. Given a catenoid defined such that

$$\begin{aligned}
x &= \cosh u_2 \cos v_2,\\
y &= \cosh u_2 \sin v_2,\\
z &= u_2,
\end{aligned}$$

show that $(E,F,G) = \left(\cosh^2 u_2, 0, \cosh^2 u_2\right)$.

20.8. Recall that the Enneper surface is given by

$$\begin{aligned}
x &= u - \frac{u^3}{3} + uv^2\\
y &= v - \frac{v^3}{3} + u^2 v\\
z &= u^2 - v^2.
\end{aligned}$$

(i) Show that the first fundamental form is given by

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

(ii) Show that the second fundamental form is given by

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

(iii) Show that the principal curvatures are

$$\kappa_1 = \frac{2}{(1+u^2+v^2)^2}, \quad \kappa_2 = -\frac{2}{(1+u^2+v^2)^2}.$$

(iv) Show that the lines of curvature are the coordinate curves.
(v) Show that the asymptotic curves are the curves of the form $u+v=C$, $u-v=C$, for some constant C.

20.9. Show that at a hyperbolic point, the principal directions bisect the asymptotic directions.

20.10. Given a pseudosphere defined such that

$$x = \frac{2\cos v}{e^u + e^{-u}},$$
$$y = \frac{2\sin v}{e^u + e^{-u}},$$
$$z = u - \frac{e^u - e^{-u}}{e^u + e^{-u}},$$

show that $K = -1$.

20.11. Prove that a general ellipsoid of equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

(a, b, c pairwise distinct) has four umbilics.

20.12. Prove that the Gaussian curvature at a point (x,y,x) of an ellipsoid of equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

has the expression

$$K = \frac{p^4}{a^2b^2c^2},$$

where p is the distance from the origin $(0,0,0)$ to the tangent plane at the point (x,y,z).

20.13. Show that the helicoid, the catenoid, and the Enneper surface are minimal surfaces, i.e., $H = 0$.

20.14. Consider two parabolas P_1 and P_2 in two orthogonal planes and such that each one passes through the focal point of the other. Given any two points $q_1 \in P_1$

and $q_2 \in P_2$, let H_{q_1,q_2} be the bisector plane of (q_1, q_2), i.e., the plane orthogonal to (q_1, q_2) and passing through the midpoint of (q_1, q_2). Prove that the envelope of the planes H_{q_1,q_2} when q_1 and q_2 vary on the parabolas P_1 and P_2 is the Enneper surface (i.e., the Enneper surface is the surface to which each H_{q_1,q_2} is tangential).

20.15. Show that if a curve on a surface S is both a line of curvature and a geodesic, then it is a planar curve.

20.16. Given a regular surface X, a *parallel surface to* X is a surface Y defined such that

$$Y(u,v) = X(u,v) + a\mathbf{N}_{(u,v)},$$

where $a \in \mathbb{R}$ is a given constant.

(i) Prove that

$$Y_u \times Y_v = (1 - 2Ha + Ka^2)(X_u \times X_v),$$

where K is the Gaussian curvature of X and H is the mean curvature of X.

(ii) Prove that the Gaussian curvature of Y is

$$\frac{K}{1 - 2Ha + Ka^2}$$

and the mean curvature of Y is

$$\frac{H - Ka}{1 - 2Ha + Ka^2},$$

where K is the Gaussian curvature of X and H is the mean curvature of X.

(iii) Assume that X has constant mean curvature $c \neq 0$. If $K \neq 0$, prove that the parallel surface Y corresponding to $a = 1/(2c)$ has constant Gaussian curvature equal to $4c^2$. Prove that the parallel surface Y corresponding to $a = 1/(2c)$ is regular except at points where $K = 0$.

(iv) Again, assume that X has constant mean curvature $c \neq 0$ and is not contained in a sphere. Show that there is a unique value of a such that the parallel surface Y has constant mean curvature $-c$. Furthermore, this parallel surface is regular at (u,v) iff $X(u,v)$ is not an umbilical point, and the Gaussian curvature of Y at (u,v) has the opposite sign to that of X.

Remark: The above results are due to Ossian Bonnet.

20.17. Given a torus of revolution defined such that

$$\begin{aligned} x &= (a + b\cos\varphi)\cos\theta, \\ y &= (a + b\cos\varphi)\sin\theta, \\ z &= b\sin\varphi, \end{aligned}$$

prove that the Gaussian curvature is given by

$$K = \frac{\cos\varphi}{b(a + b\cos\varphi)}.$$

Show that the mean curvature is given by

$$H = \frac{a + 2b\cos\varphi}{2b(a + b\cos\varphi)}.$$

20.18. (i) Given a surface of revolution defined such that

$$\begin{aligned} x &= f(v)\cos u, \\ y &= f(v)\sin u, \\ z &= g(v), \end{aligned}$$

show that the first fundamental form is given by

$$E = f(v)^2, \quad F = 0, \quad G = f'(v)^2 + g'(v)^2.$$

(ii) Show that the Christoffel symbols are given by

$$\begin{aligned} &\Gamma_{11}^1 = 0,\ \Gamma_{11}^2 = -\frac{ff'}{(f')^2 + (g')^2},\ \Gamma_{12}^1 = \frac{ff'}{f^2}, \\ &\Gamma_{12}^2 = 0,\ \Gamma_{22}^1 = 0, \qquad\qquad\qquad \Gamma_{22}^2 = \frac{f'f'' + g'g''}{(f')^2 + (g')^2}. \end{aligned}$$

(iii) Show that the equations of the geodesics are

$$\begin{aligned} u'' + \frac{2ff'}{f^2}\, u'v' &= 0, \\ v'' - \frac{ff'}{(f')^2 + (g')^2}\,(u')^2 + \frac{f'f'' + g'g''}{(f')^2 + (g')^2}\,(v')^2 &= 0. \end{aligned}$$

Show that the meridians parametrized by arc length are geodesics. Show that a parallel is a geodesic iff it is generated by the rotation of a point of the generating curve where the tangent is parallel to the axis of rotation.

(iv) Show that the first equation of geodesics is equivalent to

$$f^2 u' = c,$$

for some constant c. Since the angle θ, $0 \leq \theta \leq \pi/2$, of a geodesic with a parallel that intersects it is given by

$$\cos\theta = \frac{|X_u \cdot (X_u u' + X_v v')|}{\|X_u\|} = |fu'|,$$

and since $f = r$ is the radius of the parallel at the intersection, show that

$$r\cos\theta = c$$

for some constant $c > 0$. The equation $r\cos\theta = c$ is known as *Clairaut's relation*.

20.19. (i) Given a surface of revolution defined such that

$$\begin{aligned} x &= f(v)\cos u, \\ y &= f(v)\sin u, \\ z &= g(v), \end{aligned}$$

show that the second fundamental form is given by

$$L = -fg', \quad L = 0, \quad M = g'f'' - g''f'.$$

Conclude that the parallels and the meridians are lines of curvature.

(ii) Recall from Problem 20.18 that the first fundamental form is given by

$$E = f(v)^2, \quad F = 0, \quad G = f'(v)^2 + g'(v)^2.$$

Show that the Gaussian curvature is given by

$$K = -\frac{g'(g'f'' - g''f')}{f}.$$

Show that the parabolic points are the points where the tangent to the generating curve is perpendicular to the axis of rotation, or the points of the generating curve where the curvature is zero.

If we assume that $G = 1$, which is the case if the generating curve is parametrized by arc length, show that

$$K = -\frac{f''}{f}.$$

(iii) Show that the principal curvatures are given by

$$\kappa_1 = \frac{L}{E} = \frac{-g'}{f}, \quad \kappa_2 = \frac{N}{G} = g'f'' - g''f'.$$

20.20. (i) Is it true that if a principal curve is a plane curve, then it is a geodesic?

(ii) Is it true that if a geodesic is a plane curve, then it is a principal curve?

20.21. If X is a surface with negative Gaussian curvature, show that the asymptotic curves have the property that the torsion at (u,v) is equal to $\pm\sqrt{-K_{(u,v)}}$.

20.22. Show that if all the geodesics of a connected surface are planar curves, then this surface is contained in a plane or a sphere.

20.23. A *ruled surface* $X(t,v)$ is defined by a pair $(\alpha(t), w(t))$, where $\alpha(t)$ is some regular curve and $w(t)$ is some nonnull C^1-continuous vector in $\mathbb{R}^3$, both defined over some open interval I, with

$$X(t,v) = \alpha(t) + vw(t).$$

In other words, X consists of the one-parameter family of lines $\langle\alpha(t), w(t)\rangle$, called *rulings*. Without loss of generality, we can assume that $\|w(t)\| = 1$. In this problem we will also assume that $w'(t) \neq 0$ for all $t \in I$, in which case we say that X is *noncylindrical*.

(i) Consider the ruled surface defined such that α is the unit circle in the xy-plane, and

$$w(t) = \alpha'(t) + e_3,$$

where $e_3 = (0, 0, 1)$. Show that X can be parametrized as

$$X(t, v) = (\cos t - v \sin t,\, \sin t + v \cos t,\, v).$$

Show that X is the quadric of equation

$$x^2 + y^2 - z^2 = 1.$$

What happens if we choose $w(t) = -\alpha'(t) + e_3$?

(ii) Prove that there is a curve $\beta(t)$ on X (called the *line of striction of* X) such that

$$\beta(t) = \alpha(t) + u(t)w(t) \quad \text{and} \quad \beta'(t) \cdot w'(t) = 0$$

for all $t \in I$, for some function $u(t)$.
Hint. Show that $u(t)$ is uniquely defined by

$$u = -\frac{\alpha' \cdot w'}{w' \cdot w'}.$$

Prove that β depends only on the surface X in the following sense: If α_1 and α_2 are two curves such that

$$\alpha_2(t) + v w(t) = \alpha_1(t) + \delta(v) w(t)$$

for all $t \in I$ and all $v \in \mathbb{R}$ for some C^3-function δ, and β_1, β_2 are the corresponding lines of striction, then $\beta_1 = \beta_2$.

(iii) Writing $X(t, v)$ as

$$X(t, v) = \beta(t) + v w(t),$$

show that there is some function $\lambda(t)$ such that $\beta' \times w = \lambda w'$ and

$$\|X_t \times X_v\|^2 = (\lambda^2 + v^2)\|w'\|^2.$$

Furthermore, show that

$$\lambda = \frac{(\beta', w, w')}{\|w'\|^2}.$$

Show that the singular points (if any) occur along the line of striction $v = 0$, and that they occur iff $\lambda(t) = 0$.

(iv) Show that

$$M = \frac{(\beta', w, w')}{\|X_t \times X_v\|}, \quad N = 0,$$

and

$$K = -\frac{\lambda^2}{(\lambda^2 + v^2)^2}.$$

Conclude that the Gaussian curvature of a ruled surface satisfies $K \leq 0$, and that $K = 0$ only along those rulings that meet the line of striction at a singular point.

20.24. As in Problem 20.23, let X be a *ruled surface* $X(t,v)$ where

$$X(t,v) = \alpha(t) + vw(t)$$

and with $\|w(t)\| = 1$. In this problem we will assume that

$$(w, w', \alpha') = 0,$$

and we call such a ruled surface *developable*.

(i) Show that

$$M = \frac{(\alpha', w, w')}{\|X_t \times X_v\|}, \quad N = 0.$$

Conclude that $M = 0$, and thus that $K = 0$.

(ii) If $w(t) \times w'(t) = 0$ for all $t \in I$, show that $w(t)$ is constant and that the surface is a cylinder over a plane curve obtained by intersecting the cylinder with a plane normal to w.

If $w(t) \times w'(t) \neq 0$ for all $t \in I$, then $w'(t) \neq 0$ for all $t \in I$. Using Problem 20.23, there is a line of striction β and a function $\lambda(t)$. Check that $\lambda = 0$. If $\beta'(t) \neq 0$ for all $t \in I$, then show that the ruled surface is the tangent surface of β. If $\beta'(t) = 0$ for all $t \in I$, then show that the ruled surface is a cone.

20.25. (i) Let $\alpha \colon I \to \mathbb{R}^3$ be a curve on a regular surface S, and consider the ruled surface X defined such that

$$X(u,v) = \alpha(u) + v\mathbf{N}_{(u(t),v(t))},$$

where $\mathbf{N}_{(u(t),v(t))}$ is the unit normal vector to S at $\alpha(t)$. Prove that α is a line of curvature on S iff X is developable.

(ii) Let X be a regular surface without parabolic, planar, or umbilical points. Consider the two surfaces Y and Z (called *focal surfaces of* X, or *caustics of* X) defined such that

$$Y(u,v) = X(u,v) + \frac{1}{\kappa_1}\mathbf{N}_{(u,v)},$$
$$Z(u,v) = X(u,v) + \frac{1}{\kappa_2}\mathbf{N}_{(u,v)},$$

where κ_1 and κ_2 are the principal curvatures at (u,v).

Prove that if $(\kappa_1)_u$ and $(\kappa_2)_v$ are nowhere zero, then Y and Z are regular surfaces.

(iii) Show that the focal surfaces Y and Z are generated by the lines of strictions of the developable surfaces generated by the normals to the lines of curvatures on X. This means that if we consider the two orthogonal families $\mathscr{F}_1$ and $\mathscr{F}_2$ of lines of curvatures on X, for any line of curvature C in $\mathscr{F}_1$ (or in $\mathscr{F}_2$), the line of striction of the developable surface generated by the normals to the points of C lies on Y (or Z), and when $C \in \mathscr{F}_1$ varies on X, the corresponding line of striction sweeps Y (or Z).

What are the positions of the focal surfaces with respect to X, depending on the sign of the Gaussian curvature K? Is it possible for Y and Z to be reduced to a single point? Is it possible for Y and Z to be reduced to a curve? If Y reduces to a curve, show that X is the envelope of a one-parameter family of spheres.

20.26. Given a nonplanar regular curve f in $\mathbb{E}^3$, the surface F generated by the tangents lines to f is called the *tangent surface* of f. The tangent surface F of f may be defined by the equation

$$F(t,v) = f(t) + v\mathbf{t},$$

where $\mathbf{t} = f'(t)$. Assume that f is biregular. An *involute* of f is a curve g contained in the tangent surface of f and such that g intersects orthogonally every tangent of f. Assuming that f is parametrized by arc length, this means that every involute g of f is defined by an equation of the form

$$g(s) = f(s) + v(s)\mathbf{t}(s),$$

where $v(s)$ is a C^1-function, $\mathbf{t} = f'(s)$, and where $g'(s) \cdot \mathbf{t}(s) = 0$.

(a) Prove that

$$g'(s) = \mathbf{t} + v(s)\kappa\mathbf{n} + v'(s)\mathbf{t},$$

where $\mathbf{n}$ is the principal normal vector to f at s. Prove that the equation

$$1 + v'(s) = 0$$

must hold. Conclude that

$$v(s) = C - s,$$

where C is some constant, and thus that every involute has an equation of the form

$$g(s) = f(s) + (C - s)\mathbf{t}(s).$$

Remark: There is a physical interpretation of involutes. If a thread lying on the curve is unwound so that the unwound portion of it is always held taut in the direction of the tangent to the curve while the rest of it lies on the curve, then every point of the thread generates an involute of the curve during this motion.

(b) Consider the twisted cubic defined by

$$f(t) = \left(t, \frac{t^2}{2}, \frac{t^3}{6}\right).$$

Prove that the element of arc length is

$$ds = \left(1 + \frac{t^2}{2}\right) dt.$$

Give the equation of any involute of the twisted cubic.

Extra Credit: Plot the twisted cubic in some suitable interval and some of its involutes.

20.27. Let $\Omega\colon X \to \mathbb{E}^3$ be a surface.

(a) Assume that every point of X is an umbilic. Prove that X is contained in a sphere.

Hint. If $\kappa_1 = \kappa_2 = \kappa$ for every $(u,v) \in \Omega$, then $d\mathbf{N}(w) = -\kappa w$ for all tangent vectors $w = X_u x + X_v y$, which implies that $\mathbf{N}_u = -\kappa X_u$ and $\mathbf{N}_v = -\kappa X_v$. Prove that κ does not depend on $(u,v) \in \Omega$, i.e., it is a nonnull constant. Then prove that $X + \mathbf{N}/\kappa$ is a constant vector.

(b) Assume that every point of X is a planar point. Prove that X is contained in a plane.

Hint. This time, $\kappa_1 = \kappa_2 = 0$ for every $(u,v) \in \Omega$. Prove that $\mathbf{N}$ does not depend on $(u,v) \in \Omega$, i.e., it is a constant vector, and compute $(X \cdot \mathbf{N})_u$ and $(X \cdot \mathbf{N})_v$.

References

1. Ralph Abraham and Jerrold E. Marsden. *Foundations of Mechanics*. Addison-Wesley, second edition, 1978.
2. Lars V. Ahlfors and Leo Sario. *Riemann Surfaces*. Princeton Math. Series, No. 2. Princeton University Press, 1960.
3. Richard H. Bartels, John C. Beatty, and Brian A. Barsky. *An Introduction to Splines for Use in Computer Graphics and Geometric Modelling*. Morgan Kaufmann, first edition, 1987.
4. Marcel Berger and Bernard Gostiaux. *Géométrie différentielle: variétés, courbes et surfaces*. Collection Mathématiques. Puf, second edition, 1992. English edition: Differential geometry, manifolds, curves, and surfaces, GTM No. 115, Springer-Verlag.
5. William M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, second edition, 1986.
6. Yvonne Choquet-Bruhat, Cécile DeWitt-Morette, and Margaret Dillard-Bleick. *Analysis, Manifolds, and Physics, Part I: Basics*. North-Holland, first edition, 1982.
7. Gaston Darboux. *Leçons sur la théorie générale des surfaces, Troisième Partie*. Gauthier-Villars, first edition, 1894.
8. Gaston Darboux. *Leçons sur la théorie générale des surfaces, Quatrième Partie*. Gauthier-Villars, first edition, 1896.
9. Gaston Darboux. *Leçons sur la théorie générale des surfaces, Première Partie*. Gauthier-Villars, second edition, 1914.
10. Gaston Darboux. *Leçons sur la théorie générale des surfaces, Deuxième Partie*. Gauthier-Villars, second edition, 1915.

11. Jean Dieudonné. *Abrégé d'Histoire des Mathématiques, 1700–1900*. Hermann, first edition, 1986.
12. Manfredo P. do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.
13. Manfredo P. do Carmo. *Riemannian Geometry*. Birkhäuser, second edition, 1992.
14. Manfredo P. do Carmo. *Differential Forms and Applications*. Universitext. Springer-Verlag, first edition, 1994.
15. Gerald Farin. *NURB Curves and Surfaces, from Projective Geometry to Practical Use*. AK Peters, first edition, 1995.
16. Gerald Farin. *Curves and Surfaces for CAGD*. Academic Press, fourth edition, 1998.
17. William Fulton. *Algebraic Topology, A First Course*. GTM No. 153. Springer-Verlag, first edition, 1995.
18. Jean H. Gallier. *Curves and Surfaces in Geometric Modeling: Theory and Algorithms*. Morgan Kaufmann, first edition, 1999.
19. S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*. Universitext. Springer-Verlag, second edition, 1993.
20. R.V. Gamkrelidze (Ed.). *Geometry I.* Encyclopaedia of Mathematical Sciences, Vol. 28. Springer-Verlag, first edition, 1991.
21. Claude Godbillon. *Géométrie Différentielle et Mécanique Analytique*. Collection Méthodes. Hermann, first edition, 1969.
22. André Gramain. *Topologie des Surfaces*. Collection Sup. Puf, first edition, 1971.
23. A. Gray. *Modern Differential Geometry of Curves and Surfaces*. CRC Press, second edition, 1997.
24. Victor Guillemin and Alan Pollack. *Differential Topology*. Prentice-Hall, first edition, 1974.
25. D. Hilbert and S. Cohn-Vossen. *Geometry and the Imagination*. Chelsea Publishing Co., 1952.
26. Heinz Hopf. *Differential Geometry in the Large*. LNCS, Vol. 1000. Springer-Verlag, second edition, 1989.
27. J. Hoschek and D. Lasser. *Computer-Aided Geometric Design*. AK Peters, first edition, 1993.
28. Erwin Kreyszig. *Differential Geometry*. Dover, first edition, 1991.
29. Jacques Lafontaine. *Introduction aux Variétés Différentielles*. PUG, first edition, 1996.
30. Serge Lang. *Differential and Riemannian Manifolds*. GTM No. 160. Springer-Verlag, third edition, 1995.
31. Daniel Lehmann and Carlos Sacré. *Géométrie et Topologie des Surfaces*. Puf, first edition, 1982.
32. Charles Loop. A G^1 triangular spline surface of arbitrary topological type. *Computer-Aided Geometric Design*, 11:303–330, 1994.
33. Paul Malliavin. *Géométrie Différentielle Intrinsèque*. Enseignement des Sciences, No. 14. Hermann, first edition, 1972.
34. William S. Massey. *Algebraic Topology: An Introduction*. GTM No. 56. Springer-Verlag, second edition, 1987.
35. William S. Massey. *A Basic Course in Algebraic Topology*. GTM No. 127. Springer-Verlag, first edition, 1991.
36. John W. Milnor. *Topology from the Differentiable Viewpoint*. The University Press of Virginia, second edition, 1969.
37. John W. Milnor. *Morse Theory*. Annals of Math. Series, No. 51. Princeton University Press, third edition, 1969.
38. Henry P. Moreton. *Minimum curvature variation curves, networks, and surfaces for fair free-form shape design*. PhD thesis, University of California, Berkeley, 1993.
39. Les Piegl and Wayne Tiller. *The NURBS Book*. Monograph in Visual Communications. Springer-Verlag, first edition, 1995.
40. Richard W. Sharpe. *Differential Geometry. Cartan's Generalization of Klein's Erlangen Program*. GTM No. 166. Springer-Verlag, first edition, 1997.
41. S. Sternberg. *Lectures On Differential Geometry*. AMS Chelsea, second edition, 1983.
42. J.J. Stoker. *Differential Geometry*. Wiley Classics. Wiley-Interscience, first edition, 1989.

43. Frank Warner. *Foundations of Differentiable Manifolds and Lie Groups*. GTM No. 94. Springer-Verlag, first edition, 1983.
44. William Welch. *Serious Putty: Topological Design for Variational Curves and Surfaces*. PhD thesis, Carnegie Mellon University, Pittsburgh, Pa., 1995.