1.1. More on Frenet Frames for nD Curves

Given a curve $f\colon]a,b[\to \mathbb{E}^n$ (or $f\colon [a,b]\to \mathbb{E}^n$) of class C^p, with $p\geq n$, it is interesting to consider families $(e_1(t),\ldots,e_n(t))$ of orthonormal frames. Moreover, if for every k, with $1\leq k\leq n$, the kth derivative $f^{(k)}(t)$ of the curve $f(t)$ is a linear combination of $(e_1(t),\ldots,e_k(t))$ for every $t\in]a,b[$, then such a frame plays the role of a generalized Frenet frame. This leads to the following definition:

Definition 1.1.1 Let $f\colon]a,b[\to \mathbb{E}^n$ (or $f\colon [a,b]\to \mathbb{E}^n$) be a curve of class C^p, with $p\geq n$. A family $(e_1(t),\ldots,e_n(t))$ of orthonormal frames, where each $e_i\colon]a,b[\to \mathbb{E}^n$ is C^{n-i}-continuous for $i=1,\ldots,n-1$ and e_n is C^1-continuous, is called a *moving frame along f*. Furthermore, a moving frame $(e_1(t),\ldots,e_n(t))$ along f so that for every k, with $1\leq k\leq n$, the kth derivative $f^{(k)}(t)$ of $f(t)$ is a linear combination of $(e_1(t),\ldots,e_k(t))$ for every $t\in]a,b[$, is called a *Frenet n-frame* or *Frenet frame*.

If $(e_1(t), \ldots, e_n(t))$ is a moving frame, then

$$e_i(t) \cdot e_j(t) = \delta_{i\,j} \quad \text{for all } i, j,\ 1 \leq i, j \leq n.$$

Lemma 1.1.2 *Let $f\colon]a, b[\to \mathbb{E}^n$ (or $f\colon [a, b] \to \mathbb{E}^n$) be a curve of class C^p, with $p \geq n$, so that the derivatives $f^{(1)}(t), \ldots, f^{(n-1)}(t)$ of $f(t)$ are linearly independent for all $t \in]a, b[$. Then, there is a unique Frenet n-frame $(e_1(t), \ldots, e_n(t))$ satisfying the following conditions:*

(1) The k-frames $(f^{(1)}(t), \ldots, f^{(k)}(t))$ and $(e_1(t), \ldots, e_k(t))$ have the same orientation for all k, with $1 \leq k \leq n-1$.

(2) The frame $(e_1(t), \ldots, e_n(t))$ has positive orientation.

Proof. Since $(f^{(1)}(t), \ldots, f^{(n-1)}(t))$ is linearly independent, we can use the Gram-Schmidt orthonormalization procedure (see lemma **??**) to construct $(e_1(t), \ldots, e_{n-1}(t))$ from $(f^{(1)}(t), \ldots, f^{(n-1)}(t))$. We use the generalized cross-product to define e_n, where

$$e_n = e_1 \times \cdots \times e_{n-1}.$$

From the Gram-Schmidt procedure, it is easy to check that $e_k(t)$ is C^{n-k} for $1 \leq k \leq n-1$, and since the components of e_n are certain determinants involving the components of $(e_1, \ldots, e_{n-1})$, it is also clear that e_n is C^1. $\square$

The Frenet n-frame given by Lemma 1.1.2 is called the *distinguished Frenet n-frame*. We can now prove a generalization of the Frenet-Serret formula that gives an expression of the derivatives of a moving frame in terms of the moving frame itself.

Lemma 1.1.3 *Let* $f\colon]a,b[\to \mathbb{E}^n$ *(or* $f\colon [a,b]\to \mathbb{E}^n$*) be a curve of class* C^p*, with* $p\geq n$*, so that the derivatives* $f^{(1)}(t),\ldots,f^{(n-1)}(t)$ *of* $f(t)$ *are linearly independent for all* $t\in\,]a,b[$ *. Then, for any moving frame* $(e_1(t),\ldots,e_n(t))$*, if we write* $\omega_{ij}(t)=e_i'(t)\cdot e_j(t)$*, we have*

$$e_i'(t)=\sum_{j=1}^{n}\omega_{ij}(t)e_j(t),$$

with

$$\omega_{ji}(t)=-\omega_{ij}(t),$$

and there are some functions $\alpha_i(t)$ *so that*

$$f'(t)=\sum_{i=1}^{n}\alpha_i(t)e_i(t).$$

Furthermore, if $(e_1(t), \ldots, e_n(t))$ *is the distinguished Frenet* n*-frame associated with* f*, then we also have*

$$\alpha_1(t) = \|f'(t)\|, \qquad \alpha_i(t) = 0 \quad \textit{for} \quad i \geq 2,$$

and

$$\omega_{ij}(t) = 0 \quad \textit{for} \quad j > i + 1.$$

Proof. Since $(e_1(t), \ldots, e_n(t))$ is a moving frame, it is an orthonormal basis, and thus, $f'(t)$ and $e_i'(t)$ are linear combinations of $(e_1(t), \ldots, e_n(t))$. Also, we know that

$$e_i'(t) = \sum_{j=1}^{n} (e_i'(t) \cdot e_j(t)) e_j(t),$$

and since $e_i(t) \cdot e_j(t) = \delta_{ij}$, by differentiating, if we write $\omega_{ij}(t) = e_i'(t) \cdot e_j(t)$, we get

$$\omega_{ji}(t) = -\omega_{ij}(t).$$

Now, if $(e_1(t), \ldots, e_n(t))$ is the distinguished Frenet frame, by construction, $e_i(t)$ is a linear combination of $f^{(1)}(t), \ldots, f^{(i)}(t)$, and thus $e_i'(t)$ is a linear combination of $f^{(2)}(t), \ldots, f^{(i+1)}(t)$, hence of $(e_1(t), \ldots, e_{i+1}(t))$. $\square$

In matrix form, when $(e_1(t), \ldots, e_n(t))$ is the distinguished Frenet frame, the row vector $(e_1'(t), \ldots, e_n'(t))$ can be expressed in terms of the row vector $(e_1(t), \ldots, e_n(t))$ via a skew symmetric matrix ω, as shown below:

$$(e_1'(t), \ldots, e_n'(t)) = -(e_1(t), \ldots, e_n(t))\omega(t),$$

where

$$\omega = \begin{pmatrix} 0 & \omega_{1\,2} & & & \\ -\omega_{1\,2} & 0 & \omega_{2\,3} & & \\ & -\omega_{2\,3} & 0 & \ddots & \\ & & \ddots & \ddots & \omega_{n-1\,n} \\ & & & -\omega_{n-1\,n} & 0 \end{pmatrix}.$$

The next lemma shows the effect of a reparametrization and of a rigid motion.

Lemma 1.1.4 *Let* $f\colon]a,b[\to \mathbb{E}^n$ *(or* $f\colon [a,b]\to \mathbb{E}^n$*) be a curve of class* C^p*, with* $p\geq n$*, so that the derivatives* $f^{(1)}(t),\ldots,f^{(n-1)}(t)$ *of* $f(t)$ *are linearly independent for all* $t\in\,]a,b[$ *. Let* $h\colon \mathbb{E}^n\to\mathbb{E}^n$ *be a rigid motion, and assume that the corresponding linear isometry is* R*. Let* $\widetilde{f}=h\circ f$*. The following properties hold:*

(1) For any moving frame $(e_1(t),\ldots,e_n(t))$*, the* n*-tuple* $(\widetilde{e_1}(t),\ldots,\widetilde{e_n}(t))$*, where* $\widetilde{e_i}(t)=R(e_i(t))$*, is a moving frame along* $\widetilde{f}$*, and we have*

$$\widetilde{\omega_{ij}}(t)=\omega_{ij}(t)\quad \text{and}\quad \|\widetilde{f}'(t)\|=\|f'(t)\|.$$

(2) For any orientation-preserving diffeormorphism $\rho\colon\,]c,d[\to\,]a,b[$ *(i.e.,* $\rho'(t)>0$ *for all* $t\in\,]c,d[$ *), if we write* $\widetilde{f}=f\circ\rho$*, then for any moving frame* $(e_1(t),\ldots,e_n(t))$ *on* f*, the* n*-tuple* $(\widetilde{e_1}(t),\ldots,\widetilde{e_n}(t))$*, where* $\widetilde{e_i}(t)=e_i(\rho(t))$*, is a moving frame on* $\widetilde{f}$*.*

Furthermore, if $\|\widetilde{f}'(t)\| \neq 0$*, then*

$$\frac{\widetilde{\omega_{ij}}(t)}{\|\widetilde{f}'(t)\|} = \frac{\omega_{ij}(\rho(t))}{\|f'(\rho(t))\|}.$$

The proof is straightforward and is omitted.

The above lemma suggests the definition of the curvatures $\kappa_1, \ldots, \kappa_{n-1}$.

Definition 1.1.5 Let $f\colon\,]a, b[\, \to \mathbb{E}^n$ (or $f\colon [a, b] \to \mathbb{E}^n$) be a curve of class C^p, with $p \geq n$, so that the derivatives $f^{(1)}(t), \ldots, f^{(n-1)}(t)$ of $f(t)$ are linearly independent for all $t \in\,]a, b[$. If $(e_1(t), \ldots, e_n(t))$ is the distinguished Frenet frame associated with f, we define the *ith curvature*, κ_i, *of* f, by

$$\kappa_i(t) = \frac{\omega_{i\, i+1}(t)}{\|f'(t)\|},$$

with $1 \leq i \leq n - 1$.

Observe that the matrix $\omega(t)$ can be written as

$$\omega(t) = \|f'(t)\|\kappa(t),$$

where

$$\kappa = \begin{pmatrix} 0 & \kappa_{1\,2} & & & \\ -\kappa_{1\,2} & 0 & \kappa_{2\,3} & & \\ & -\kappa_{2\,3} & 0 & \ddots & \\ & & \ddots & \ddots & \kappa_{n-1\,n} \\ & & & -\kappa_{n-1\,n} & 0 \end{pmatrix}.$$

The matrix κ is sometimes called the *Cartan matrix*.

Lemma 1.1.6 *Let $f\colon]a,b[\to \mathbb{E}^n$ (or $f\colon [a,b] \to \mathbb{E}^n$) be a curve of class C^p, with $p \geq n$, so that the derivatives $f^{(1)}(t), \ldots, f^{(n-1)}(t)$ of $f(t)$ are linearly independent for all $t \in]a,b[$. Then for every i, with $1 \leq i \leq n-2$, we have $\kappa_i(t) > 0$.*

Proof. Lemma 1.1.2 shows that $e_1, \ldots, e_{n-1}$ are expressed in terms of $f^{(1)}, \ldots, f^{(n-1)}$ by a triangular matrix (a_{ij}), whose diagonal entries $a_{i\,i}$ are strictly positive, i.e., we have

$$e_i = \sum_{j=1}^{i} a_{ij} f^{(j)},$$

for $i = 1, \ldots, n-1$, and thus,

$$f^{(i)} = \sum_{j=1}^{i} b_{ij} e_j,$$

for $i = 1, \ldots, n-1$, with $b_{i\,i} = a_{i\,i}^{-1} > 0$. Then, since $e_{i+1} \cdot f^{(j)} = 0$ for $j \leq i$, we get

$$\|f'\| \kappa_i = \omega_{i\,i+1} = e_i' \cdot e_{i+1} = a_{i\,i} f^{(i+1)} \cdot e_{i+1} = a_{i\,i} b_{i+1\,i+1},$$

and since $a_{i\,i} b_{i+1\,i+1} > 0$, we get $\kappa_i > 0$ $(i = 1, \ldots, n-2)$. $\square$

We conclude by exploring to what extent the curvatures $\kappa_1, \ldots,$ κ_{n-1} determine a curve satisfying the nondegeneracy conditions of Lemma 1.1.2. Basically, such curves are defined up to a rigid motion.

Lemma 1.1.7 *Let* $f\colon]a,b[\to \mathbb{E}^n$ *and* $\widetilde{f}\colon]a,b[\to \mathbb{E}^n$ *(or* $f\colon [a,b] \to \mathbb{E}^n$ *and* $\widetilde{f}\colon [a,b] \to \mathbb{E}^n$*) be two curves of class* C^p*, with* $p \geq n$*, and satisfying the nondegeneracy conditions of Lemma 1.1.2. Denote the distinguished Frenet frames associated with* f *and* $\widetilde{f}$ *by* $(e_1(t), \ldots, e_n(t))$ *and* $(\widetilde{e_1}(t), \ldots, \widetilde{e_n}(t))$*. If* $\kappa_i(t) = \widetilde{\kappa_i}(t)$ *for every* i*, with* $1 \leq i \leq n-1$*, and* $\|f'(t)\| = \|\widetilde{f}'(t)\|$ *for all* $t \in]a,b[$ *, then there is a unique rigid motion* h *so that*

$$\widetilde{f} = h \circ f.$$

Proof. Fix $t_0 \in]a,b[$. First of all, there is a unique rigid motion h so that

$$h(f(t_0)) = \widetilde{f}(t_0) \quad \text{and} \quad R(e_i(t_0)) = \widetilde{e_i}(t_0),$$

for all i, with $1 \leq i \leq n$, where R is the linear isometry associated with h (in fact, a rotation). Consider the curve $\overline{f} = h \circ f$. The hypotheses of the lemma and Lemma 1.1.4, imply that

$$\overline{\omega_{ij}}(t) = \widetilde{\omega_{ij}}(t) = \omega_{ij}(t), \quad \|\overline{f}'(t)\| = \|\widetilde{f}'(t)\| = \|f'(t)\|,$$

and, by construction,
$(\overline{e_1}(t_0), \ldots, \overline{e_n}(t_0)) = (\widetilde{e_1}(t_0), \ldots, \widetilde{e_n}(t_0))$ and
$\overline{f}(t_0) = \widetilde{f}(t_0)$. Let

$$\delta(t) = \sum_{i=1}^{n} (\overline{e_i}(t) - \widetilde{e_i}(t)) \cdot (\overline{e_i}(t) - \widetilde{e_i}(t)).$$

Then, we have

$$
\begin{aligned}
\delta'(t) &= 2\sum_{i=1}^{n}(\overline{e_i}(t) - \widetilde{e_i}(t)) \cdot (\overline{e_i}'(t) - \widetilde{e_i}'(t)) \\
&= -2\sum_{i=1}^{n}(\overline{e_i}(t) \cdot \widetilde{e_i}'(t) + \widetilde{e_i}(t) \cdot \overline{e_i}'(t)).
\end{aligned}
$$

Using the Frenet equations, we get

$$
\begin{aligned}
\delta'(t) &= -2\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}\overline{e_i} \cdot \widetilde{e_j} - 2\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}\overline{e_j} \cdot \widetilde{e_i} \\
&= -2\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}\overline{e_i} \cdot \widetilde{e_j} - 2\sum_{j=1}^{n}\sum_{i=1}^{n}\omega_{ji}\overline{e_i} \cdot \widetilde{e_j} \\
&= -2\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}\overline{e_i} \cdot \widetilde{e_j} + 2\sum_{j=1}^{n}\sum_{i=1}^{n}\omega_{ij}\overline{e_i} \cdot \widetilde{e_j} \\
&= 0,
\end{aligned}
$$

since ω is skew symmetric. Thus, $\delta(t)$ is constant, and since the Frenet frames at t_0 agree, we get $\delta(t) = 0$.

Then, $\overline{e_i}(t) = \widetilde{e_i}(t)$ for all i, and since $\|\overline{f}'(t)\| = \|\widetilde{f}'(t)\|$, we have

$$\overline{f}'(t) = \|\overline{f}'(t)\|\overline{e_1}(t) = \|\widetilde{f}'(t)\|\widetilde{e_1}(t) = \widetilde{f}'(t),$$

so that $\overline{f}(t) - \widetilde{f}(t)$ is constant. However, $\overline{f}(t_0) = \widetilde{f}(t_0)$, and so, $\overline{f}(t) = \widetilde{f}(t)$, and $\widetilde{f} = \overline{f} = h \circ f$. $\square$

Lemma 1.1.8 *Let $\kappa_1, \ldots, \kappa_{n-1}$ be functions defined on some open $]a, b[$ containing 0 with κ_i C^{n-i-1}-continuous for $i = 1, \ldots, n-1$, and with $\kappa_i(t) > 0$ for $i = 1, \ldots, n-2$ and all $t \in\,]a, b[$. Then, there is curve $f\colon\,]a, b[\, \to \mathbb{E}^n$ of class C^p, with $p \geq n$, satisfying the nondegeneracy conditions of Lemma 1.1.2, so that $\|f'(t)\| = 1$ and f has the $n-1$ curvatures $\kappa_1(t), \ldots, \kappa_{n-1}(t)$.*

Proof. Let $X(t)$ be the matrix whose columns are the vectors $e_1(t), \ldots, e_n(t)$ of the Frenet frame along f. Consider the system of ODE's,

$$X'(t) = -X(t)\kappa(t),$$

with initial conditions $X(0) = I$, where $\kappa(t)$ is the skew symmetric matrix of curvatures. By a standard result in ODE's, there is a unique solution $X(t)$.

We claim that $X(t)$ is an orthogonal matrix. For this, note that

$$\begin{aligned}(XX^\top)' &= X'X^\top + X(X^\top)' = -X\kappa X^\top - X\kappa^\top X^\top \\ &= -X\kappa X^\top + X\kappa X^\top = 0.\end{aligned}$$

Since $X(0) = I$, we get $XX^\top = I$. If $F(t)$ is the first column of $X(t)$, we define the curve f by

$$f(s) = \int_0^s F(t)dt,$$

with $s \in]a, b[$. It is easily checked that f is a curve parametrized by arc length, with Frenet frame $X(s)$, and with curvatures κ_i's. $\square$

Chapter 2

Basics of The Differential Geometry of Surfaces

2.1. Introduction

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.

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 parameterized locally.

Thus, it is important and useful to study parameterized patches.

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 in immersed in the ambiant space, whereas extrinsic properties depend on properties of the ambiant 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.

In this chapter, we focus exclusively on the study of local properties.

By studying the properties of the curvature of curves on a surface, we will be led to the first and to the second fundamental form of a surface.

The study of the normal and of the 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 form.

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.

2.2. Parameterized 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.

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

A surface 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.

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

$$\left(\frac{du}{dt}(t), \frac{dv}{dt}(t)\right) \neq (0, 0)$$

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*.

The tangent vector $\frac{dC}{dt}(t)$ 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

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

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

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

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

are denoted by $\dot{C}(t)$, $\dot{u}(t)$ and $\dot{v}(t)$, or even as $\dot{C}$, $\dot{u}$, and $\dot{v}$.

When the curve C is parameterized by arc length s, we denote

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

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

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

Using these notations, $\dot{C}(t)$ is expressed as follows:

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

or simply as

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

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 2.2.1 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* $(u, v) \in \Omega$ iff $X_u \times X_v \neq \overrightarrow{0}$, and we also say that $p = X(u, v)$ is *a regular point of* X. If $X_u \times X_v = \overrightarrow{0}$, we say that $p = X(u, v)$ is a *singular point of* X.

The surface X is *regular on* Ω iff $X_u \times X_v \neq \overrightarrow{0}$, 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 as $T_p(X)$.

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 $AT_p(X)$?).

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 $\mathbf{N}_p$.

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. For example, considering the surface $X = (x(u, v), y(u, v), z(u, v))$ defined such that

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

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,

$$
\begin{aligned}
X_r \times X_s &= \left(X_u \frac{\partial u}{\partial r} + X_v \frac{\partial v}{\partial r}\right) \times \left(X_u \frac{\partial u}{\partial s} + X_v \frac{\partial v}{\partial s}\right) \\
&= \left(\frac{\partial u}{\partial r}\frac{\partial v}{\partial s} - \frac{\partial u}{\partial s}\frac{\partial v}{\partial r}\right) X_u \times X_v \\
&= \det \begin{pmatrix} \dfrac{\partial u}{\partial r} & \dfrac{\partial u}{\partial s} \\[2ex] \dfrac{\partial v}{\partial r} & \dfrac{\partial v}{\partial s} \end{pmatrix} X_u \times X_v \\
&= \frac{\partial(u, v)}{\partial(r, s)} X_u \times X_v,
\end{aligned}
$$

denoting the Jacobian determinant of the map $(r,s)\mapsto(u,v)$ as $\frac{\partial(u,v)}{\partial(r,s)}$.

Then, the relationship between the unit vectors $\mathbf{N}_{(u,v)}$ and $\mathbf{N}_{(r,s)}$ is

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

We will therefore restrict our attention to changes of variables such that the Jacobian determinant $\frac{\partial(u,v)}{\partial(r,s)}$ 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

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

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.

2.3. 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 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 $\|\dot{C}\|^2$, it is positive definite.

This quadratric form plays a major role in the theory of surfaces, and deserves an official definition.

Definition 2.3.1 Given a 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 as 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 as $\langle (x_1, y_1), (x_2, y_2) \rangle_p$.

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

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

Thus, the length $l(pq)$ of an arc of curve on the surface joining $p = X(u(t_0), v(t_0))$ 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 $ds^2 = E\,du^2 + 2F\,dudv + G\,dv^2$ as a *Riemannian metric*. The symmetric matrix associated with the first fundamental form is also denoted as

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

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 assume X injective.

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

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

We just discovered that, contrary 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.

For example, 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 of the catenoid.

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

$$x = u_1 \cos v_1,$$
$$y = u_1 \sin v_1,$$
$$z = v_1.$$

This is the surface generated by a line parallel to the xOy plane, touching the z axis, and also touching an helix of axis Oz.

It is easily verified that $(E, F, G) = (1, 0, u_1^2 + 1)$. The figure below shows a portion of helicoid corresponding to $0 \leq v_1 \leq 2\pi$ and $-2 \leq u_1 \leq 2$.

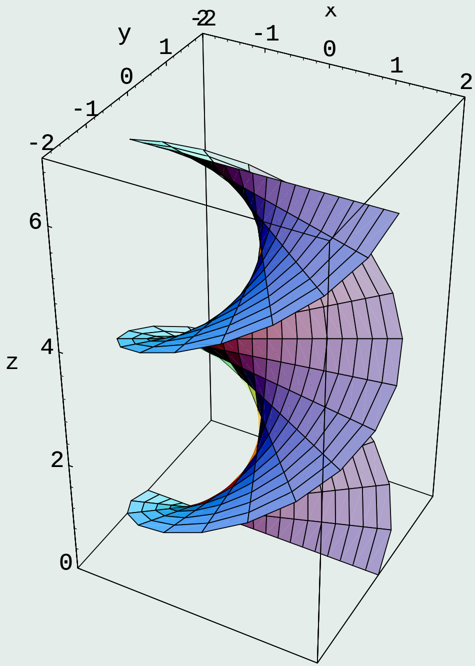

Figure 2.1: An helicoid

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.

It is easily verified that

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

The figure below shows a portion of catenoid corresponding to $0 \leq v_2 \leq 2\pi$ and $-2 \leq u_2 \leq 2$.

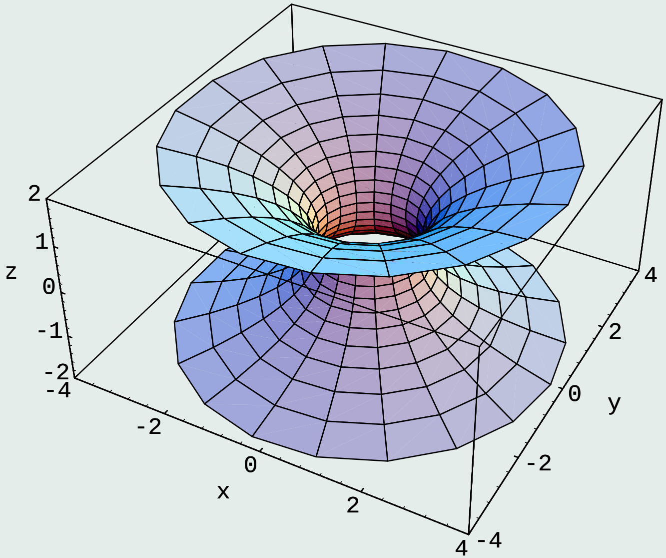

Figure 2.2: A catenoid

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 parameterization 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 made a "slit" in the catenoid (a portion of meridian), and thus we can open up the catenoid and deform it into the helicoid).

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

2.4. 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 parameterized by arc length, we will see that the vector $X''(s)$ (which is equal to $\kappa\overrightarrow{n}$, where $\overrightarrow{n}$ is the principal normal to the curve C at p, and κ is the curvature) can be written as

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

where $\mathbf{N}$ is the normal to the surface at p, and $\kappa_g \overrightarrow{n_g}$ 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 only depends 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 a least C^3-continuous, and assume that the curve is parameterized by arc length.

We saw in Chapter **??**, section **??**, 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 $(\overrightarrow{t}, \overrightarrow{n}, \overrightarrow{b})$ called the Frenet frame, where

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

The vector $\overrightarrow{t}$ is the unit *tangent vector*, the vector $\overrightarrow{n}$ is called the *principal normal*, and the vector $\overrightarrow{b}$ is called the *binormal*.

Furthermore the curvature κ at s is $\kappa = \|f''(s)\|$, and thus,

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

The principal normal $\overrightarrow{n}$ 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 parameterized by arc length, the vector $f'(s)$ is a unit vector, and thus

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

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 parameterized by arc length, which implies that

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

we have

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

and $\overrightarrow{t} = X_u u' + X_v v'$ is indeed a unit tangent vector to the curve and to the surface, but $\overrightarrow{n}$ 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 $(\overrightarrow{t}, \overrightarrow{n}, \overrightarrow{b})$ associated with the curve C is not really adequate, since both $\overrightarrow{n}$ and $\overrightarrow{b}$ will vary with C (and $\overrightarrow{n}$ 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 $(\overrightarrow{t}, \overrightarrow{n_g}, \mathbf{N})$ defined as follows.:

Definition 2.4.1 Given a surface X, given any curve $C\colon t \mapsto X(u(t), v(t))$ on X, for any point p on X, the orthonormal frame $(\overrightarrow{t}, \overrightarrow{n_g}, \mathbf{N})$ is defined such that

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

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

For simplicity of notation, we will often drop arrows above vectors if no confusion may arise.

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

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

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

The component $\kappa_N\mathbf{N}$ is the orthogonal projection of $\kappa\overrightarrow{n}$ onto the normal direction $\mathbf{N}$, and for this reason κ_N is called the *normal curvature of C at p*.

The component $\kappa_g\overrightarrow{n_g}$ is the orthogonal projection of $\kappa\overrightarrow{n}$ 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 $X'' = \kappa \overrightarrow{n}$ into its normal component (along $\mathbf{N}$) and its tangential component, we use a neat trick suggested by Eugenio Calabi.

Recall that

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

Using this identity, we have

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

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

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

we can write

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

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

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

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 determinant of the three vectors.

Some authors (including Gauss himself and Darboux) use the notation

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

and we also have

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

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 2.4.2 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 as II_p. For a curve C on the surface X (parameterized 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 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, etc.).

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 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 2.4.3 *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 can easily give an expression for κ_N for an arbitrary parameterization.

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

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

and from

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

we get

$$\kappa_N = \frac{L\dot{u}^2 + 2M\dot{u}\dot{v} + N\dot{v}^2}{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}.$$

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

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 $\overrightarrow{t} = X'$, since the component $\kappa_N \mathbf{N} \cdot \overrightarrow{t}$ is zero, we have $X_t'' \cdot \overrightarrow{t} = 0$, and thus X_t'' is also orthogonal to $\overrightarrow{t}$, which means that it is collinear with $\overrightarrow{n_g} = \mathbf{N} \times \overrightarrow{t}$.

Therefore, we showed that

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

where

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

and

$$\kappa_g \overrightarrow{n_g} = (\mathbf{N} \times (X_{uu}(u')^2 + 2X_{uv}u'v' + X_{vv}(v')^2)) \times \mathbf{N} \\ + X_u u'' + X_v v''.$$

The term $\kappa_g \overrightarrow{n_g}$ is worth an official definition.

Definition 2.4.4 Given a surface X, given any curve $C\colon t \mapsto X(u(t), v(t))$ on X, for any point p on X, the quantity κ_g appearing in the expression

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

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 $\kappa_g \overrightarrow{n_g}$ in terms of the basis (X_u, X_v).

2.5. 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 $\kappa_g \overrightarrow{n_g}$.

We now show that κ_n 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 $\overrightarrow{n_g}$ is in the tangent space $T_p(X)$, and since (X_u, X_v) is a basis of $T_p(X)$, we can write

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

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

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

However,

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

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

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

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:

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

On the other hand,

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

Dotting with X_u and X_v, we get

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

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 $[u_\alpha\,v_\beta;\,u_\gamma]$ as $[\alpha\,\beta;\,\gamma]$.

Doing so, and remembering that

$$\kappa \overrightarrow{n} \cdot X_u = EA + FB,$$
$$\kappa \overrightarrow{n} \cdot X_v = FA + GB,$$

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

$$A = u_1'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma^1_{i\,j}\, u_i' u_j',$$

$$B = u_2'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma^2_{i\,j}\, u_i' u_j',$$

and

$$\kappa_g \overrightarrow{n_g} =$$

$$\left(u_1'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma^1_{i\,j}\, u_i' u_j' \right) X_u + \left(u_2'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma^2_{i\,j}\, u_i' u_j' \right) X_v.$$

We summarize all the above in the following lemma.

Lemma 2.5.1 *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*

$$\kappa_g \overrightarrow{n_g} =$$

$$\left(u_1'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma^1_{i\,j}\, u_i' u_j'\right) X_u + \left(u_2'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma^2_{i\,j}\, u_i' u_j'\right) X_v,$$

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

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

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

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

Note that

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

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 only depend on the first fundamental form, but in fact they do!

After some calculations, we have the following formulae showing that the Christoffel symbols only depend 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 to generalize to Riemannian manifolds. In the next section, we study the variation of the normal curvature.

2.6. 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 form. We also classify points on a surface according to the value and sign of the Gaussian curvature.

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

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 $\overrightarrow{t} = X_u x + X_v y$, the normal curvature is given by

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

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 $(\overrightarrow{e_1}, \overrightarrow{e_2})$ defined such that

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

is indeed an orthonormal frame.

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

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

where $w = \sqrt{EG - F^2}$.

We can now compute $\kappa_N(\overrightarrow{t})$, and we get

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

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

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

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

where

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

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 $\theta_0 + \frac{\pi}{2}$ and $\theta_0 + \frac{3\pi}{2}$, where $\cos 2\theta_0 = \frac{A}{C}$ and $\sin 2\theta_0 = \frac{B}{C}$.

The curvatures κ_1 and κ_2 play a major role in surface theory.

Definition 2.6.1 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 $H = \frac{\kappa_1 + \kappa_2}{2}$ 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 $\theta_0 + \frac{\pi}{2}$ corresponding to κ_1 and κ_2 are orthogonal. Note that

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

After some laborious calculations, we get the following (famous) formulae for the mean curvature and the Gaussian curvature:

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

We showed 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 $(\overrightarrow{e_1}, \overrightarrow{e_2})$.

We also showed that the angle θ_0 such that $\cos 2\theta_0 = \frac{A}{C}$ and $\sin 2\theta_0 = \frac{B}{C}$, plays a special role.

Indeed, it determines one of the principal directions.

If we rotate the basis $(\overrightarrow{e_1}, \overrightarrow{e_2})$ and pick a frame $(\overrightarrow{f_1}, \overrightarrow{f_2})$ 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 get

$$\kappa_N(\theta) = \kappa_1 \cos^2\varphi + \kappa_2 \sin^2\varphi.$$

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

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

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).

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

Definition 2.6.2 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 don't.

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! On a monkey saddle, there is a planar point. The principal directions at that point are undefined.

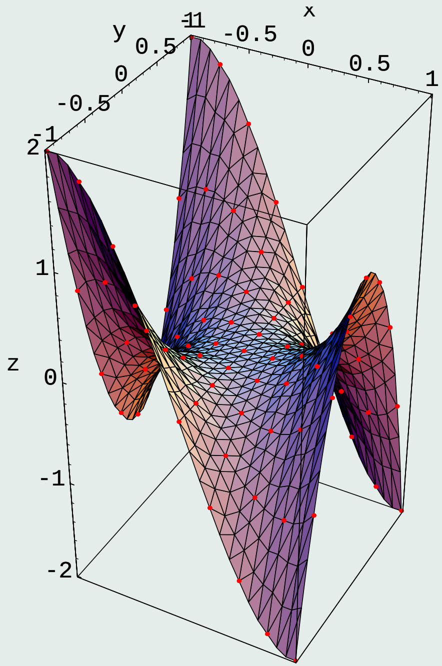

Figure 2.3: A monkey saddle

For an umbilical point, we have $\kappa_1 = \kappa_2 \neq 0$.

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

This can only happen 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.

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.

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 on the following picture).

The elliptic points are on the outside part of the torus (with normal facing outward), delimited by the two parabolic circles.

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

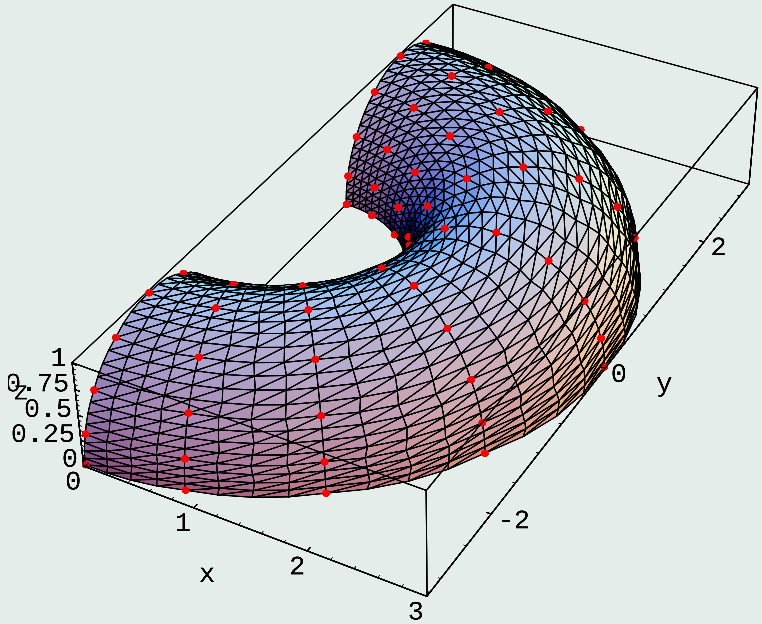

Figure 2.4: Portion of torus

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 only happen 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 $L\left(\frac{x}{y}\right)^2 + 2M\frac{x}{y} + N = 0$ 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 $K = \frac{1}{R^2}$.

Perhaps surprisingly, there are other surfaces of constant positive curvature besides the sphere.

There are surfaces of constant negative curvature, say $K = -1$. A famous one is the *pseudosphere*, 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 parameterization is given by:

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

over $]0, 2\pi[\, \times \mathbb{R}$.

The pseudosphere has a circle of singular points (for $u = 0$). The figure below shows a portion of pseudosphere.

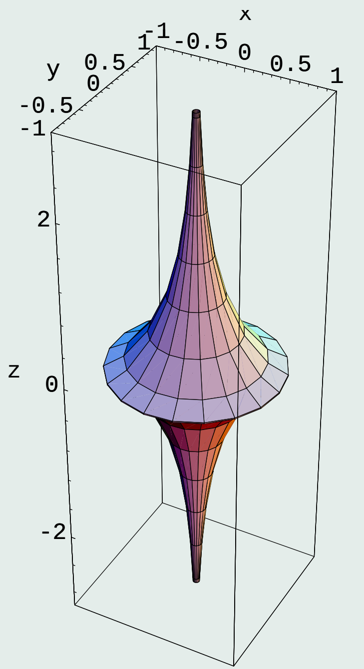

Figure 2.5: A pseudosphere

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

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

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

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

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.

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 shorlines 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.

2.7. The Gauss Map and its Derivative $d\mathbf{N}$

Given a surface $X\colon \Omega \to \mathbb{E}^3$, given 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 as $\mathbf{N}\colon X \to S^2$.

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 $\mathcal{S}_p = -d\mathbf{N}_p$, often called the *shape operator*.

Then, it is easily shown that the second fundamental form $\mathrm{II}_p(\overrightarrow{t})$ can be expressed as

$$\mathrm{II}_p(\overrightarrow{t}) = \langle \mathcal{S}_p(\overrightarrow{t}), \overrightarrow{t} \rangle_p,$$

where $\langle -, - \rangle$ is the inner product associated with the first fundamental form.

Thus, the Gaussian curvature is equal to the determinant of $\mathcal{S}_p$, 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 only depends 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.

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 as $\mathbf{N}\colon \Omega \to S^2$.

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

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

Lemma 2.7.1 *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

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

The equations

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

are know as the *Weingarten equations* (in matrix form).

If we recall from Section 2.6 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.

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

This is recorded in the following lemma that also shows that the eigenvectors of $J(d\mathbf{N}_{(u,v)})$ correspond to the principal directions:

Lemma 2.7.2 *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 $-\frac{1}{2}\mathrm{trace}(J(d\mathbf{N}_{(u,v)}))$.*

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 $\mathcal{S}_{(u,v)} = -d\mathbf{N}_{(u,v)}$ instead of $d\mathbf{N}_{(u,v)}$.

The map $\mathcal{S}_{(u,v)}$ 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 2.7.3 *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*

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

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

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

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

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 $\mathcal{S}_{(u,v)}$ is self-adjoint with respect to I can be used to give a fancier proof of the fact that $\mathcal{S}_{(u,v)}$ 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 [**?**]. Our proof is more basic and from first principles.

2.8. The Dupin Indicatrix

The second fundamental form shows up again when we study the deviation of a surface from its tangent plane in the 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).

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

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

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

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

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

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 $\rho(u, v) = \pm\frac{1}{2}\epsilon^2$ will look like portions of the curves of equation

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

Letting $u - u_0 = \epsilon x$ and $v - v_0 = \epsilon y$, 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 $\overrightarrow{e_1}$ and $\overrightarrow{e_2}$ 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

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

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

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

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

$$\frac{x^2}{\sqrt{\frac{1}{\kappa_1}}} - \frac{y^2}{\sqrt{-\frac{1}{\kappa_2}}} = 1,$$

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

$$-\frac{x^2}{\sqrt{-\frac{1}{\kappa_1}}} + \frac{y^2}{\sqrt{\frac{1}{\kappa_2}}} = 1,$$

if $\kappa_1 < 0$ and $\kappa_2 > 0$.

These hyperbolae share the same asymptotes, which are the asymptotic directions as defined in Section 2.6, 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 [**?**], Chapter IV, page 192, is wrong!

Therfore, 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).

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

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).

2.9. The *Theorema Egregium* of Gauss, the Equations of Codazzi-Mainardi, and Bonnet's Theorem

In Section 2.5, we expressed the geodesic curvature in terms of the Christoffel symbols, and we also showed that these symbols only depend on E, F, G, i.e., on the first fundamental form.

In Section 2.7, we expressed $\mathbf{N}_u$ and $\mathbf{N}_v$ in terms of the coefficients of the first and the second fundamental form.

At first glance, given any six functions E, F, G, L, M, N which 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).

Remarkably, these compatibility equations are just what it takes to insure 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

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

and since

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

we get the equations (due to Gauss):

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

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

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

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 2.7 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

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

we see that the partial derivatives of X_u, X_v and $\mathbf{N}$ can be expressed in terms of the coefficient 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 obain 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 2.9.1 *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 [**?**] (Volume III, page 246), a way of proving theorem 2.9.1 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

$$\begin{aligned} D &= (X_u, X_v, X_{uu}), \\ D' &= (X_u, X_v, X_{uv}), \\ D'' &= (X_u, X_v, X_{vv}). \end{aligned}$$

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:

$$
\begin{aligned}
M_u - L_v &= \Gamma^2_{1\,1} N - (\Gamma^2_{1\,2} - \Gamma^1_{1\,1}) M - \Gamma^1_{1\,2} L, \\
N_u - M_v &= \Gamma^2_{1\,2} N - (\Gamma^2_{2\,2} - \Gamma^1_{1\,2}) M - \Gamma^1_{2\,2} L.
\end{aligned}
$$

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 2.9.2 *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, [**?**] Appendix to Chapter 4, pp. 309-314). Finally, we state Bonnet's theorem.

Theorem 2.9.3 *Let E, F, G, L, M, N be any 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 $X : \Omega \to \mathbb{E}^3$ 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.*

2.10. 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.

Definition 2.10.1 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 2.7.2 of Section 2.7 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 2.7.2:

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

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

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

From the above equation, we see that the u-lines and the v-lines are the lines of curvatures iff $F = M = 0$.

Generally, this differential equation does not have closed-form solutions.

There is another notion which is useful in understanding lines of curvature, the geodesic torsion.

Let $C: s \mapsto X(u(s), v(s))$ be a curve on X assumed to be parameterized 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 $(\overrightarrow{e_1}, \overrightarrow{e_2}, \mathbf{N})$, known as the *Darboux frame*, where $\overrightarrow{e_1}$ and $\overrightarrow{e_2}$ are unit vectors corresponding to the principal directions, $\mathbf{N}$ is the normal to the surface at $X(u(0), v(0))$, and $\mathbf{N} = \overrightarrow{e_1} \times \overrightarrow{e_2}$.

It is interesting to study the quantity $\dfrac{d\mathbf{N}_{(u,v)}}{ds}(0)$.

If $\overrightarrow{t} = C'(0)$ is the unit tangent vector at $X(u(0), v(0))$, we have another orthonormal frame considered in Section 2.4, namely $(\overrightarrow{t}, \overrightarrow{n_g}, \mathbf{N})$, where $\overrightarrow{n_g} = \mathbf{N} \times \overrightarrow{t}$, and if φ is the angle between $\overrightarrow{e_1}$ and $\overrightarrow{t}$ we have

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

Lemma 2.10.2 *Given a curve* $C\colon s \mapsto X(u(s), v(s))$ *parameterized by arc length on a surface* X, *we have*

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

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

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

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 curvatures 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 $\overrightarrow{n}$ 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 $\overrightarrow{n}$ so that $\cos\theta = \mathbf{N} \cdot \overrightarrow{n}$, we claim that

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

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

Lemma 2.10.3 *Given a curve $C\colon s \mapsto X(u(s), v(s))$ parameterized 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 $\overrightarrow{n}$ to C at $s = 0$.

Note that the geodesic torsion only depends 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 2.10.3 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 $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ of surfaces form a *triply orthogonal system* for U, if for every point $p \in U$, there is a unique surface from each family $\mathcal{F}_i$ passing through p, where $i = 1, 2, 3$, and any two of these surfaces intersect orthogonally along their curve of intersection.

Theorem 2.10.4 *The surfaces of a triply orthogonal system intersect each other along lines of curvature.*

A nice application of theorem 2.10.4 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 [?], Chapter 10, Sections 10.2.2.3, 10.4.9.5, and 10.6.8.3, and Hilbert and Cohn-Vossen [?], Chapter 4, Section 28).

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

Definition 2.10.5 Given a surface X, an *asymptotic line* is a curve $C\colon t \mapsto X(u(t), v(t))$ 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 $F \neq 0$).

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 curvatures and the asymptotic lines.

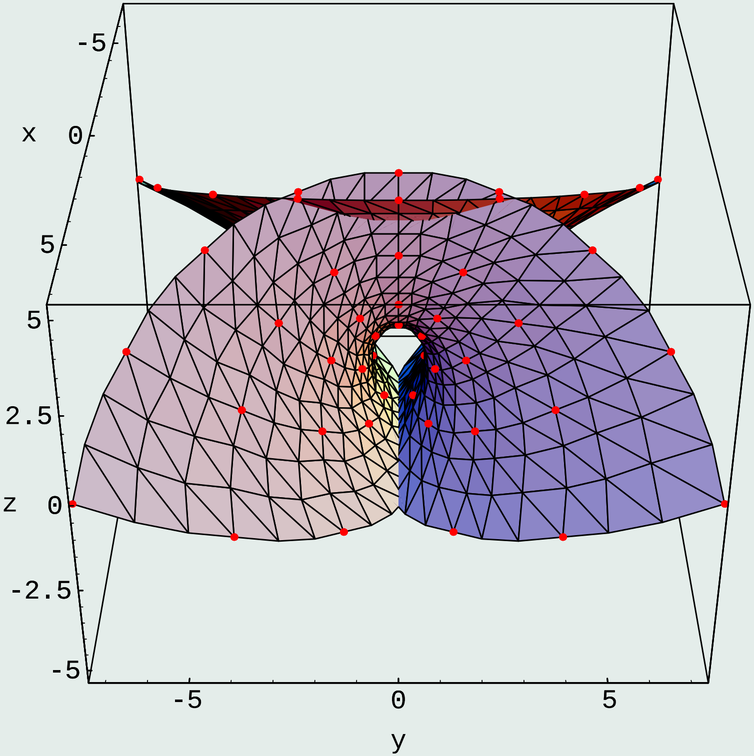

Figure 2.6: the Enneper surface

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 [**?**], Chapter IV, Section 29, page 197.

We now turn briefly to geodesics.

2.11. 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 only true *locally, not globally*.

More precisely, given a surface X, given 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 $\overrightarrow{n}$ must be parallel to the normal $\mathbf{N}$ to the surface along the curve segment from p to q.

If C is parameterized 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 [**?**] or Berger and Gostiaux [**?**]), but for simplicity, we stick to the definition in terms of the geodesic curvature.

Definition 2.11.1 Given a surface $X: \Omega \to \mathbb{E}^3$, a *geodesic line, or geodesic,* is a regular curve $C: I \to \mathbb{E}^3$ 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 $\overrightarrow{n}$ 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 $\overrightarrow{n}$ 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 2.5.1.

Since the tangential part of the curvature at a point is given by

$$\kappa_g \overrightarrow{n_g} = \left(u_1'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma^1_{i\,j}\, u_i' u_j' \right) X_u + \left(u_2'' + \sum_{\substack{i=1,2\\ j=1,2}} \Gamma^2_{i\,j}\, u_i' u_j' \right) X_v,$$

the differential equations for geodesics are

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

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

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

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 [?], Chapter 4, Section 4.7):

Lemma 2.11.2 *Given a surface X, for every point $p = X(u,v)$ on X, for every nonnull tangent vector $\overrightarrow{v} \in T_{(u,v)}(X)$ at p, there is some $\epsilon > 0$ and a unique curve $\gamma\colon]-\epsilon, \epsilon[\to \mathbb{E}^3$ on the surface X, such that γ is a geodesic, $\gamma(0) = p$, and $\gamma'(0) = \overrightarrow{v}$.*

To emphasize that the geodesic γ depends on the initial direction $\overrightarrow{v}$, we often write $\gamma(t, \overrightarrow{v})$ intead 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.

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 [**?**], Chapter 4, Section 4, and the problems.

The geodesics on an ellipsoid are also fascinating, see Berger and Gostiaux [**?**], Section 10.4.9.5, and Hilbert and Cohn-Vossen [**?**], 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 parameterize 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

$$\exp_p \colon B_\epsilon \to X(\Omega),$$

such that for every $\overrightarrow{v} \in B_\epsilon$ with $\overrightarrow{v} \neq \overrightarrow{0}$,

$$\exp_p(\overrightarrow{v}) = \gamma(1, \overrightarrow{v}),$$

where $\gamma(t, \overrightarrow{v})$ is the unique geodesic segment such that $\gamma(0, \overrightarrow{v}) = p$ and $\gamma'(0, \overrightarrow{v}) = \overrightarrow{v}$.

Furthermore, for B_ϵ small enough, $\exp_p$ is a diffeomorphism.

It turns out that $\exp_p(\overrightarrow{v})$ is the point q obtained by "laying off" a length equal to $\|\overrightarrow{v}\|$ along the unique geodesic that passes through p in the direction $\overrightarrow{v}$.

Lemma 2.11.3 *Given a surface $X\colon \Omega \to \mathbb{E}^3$, for every $\overrightarrow{v} \neq \overrightarrow{0}$ in $\mathbb{R}^2$, if*

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

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

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

is also a geodesic, and

$$\gamma(t, \lambda\overrightarrow{v}) = \gamma(\lambda t, \overrightarrow{v}).$$

From lemma 2.11.3, for $\overrightarrow{v} \neq \overrightarrow{0}$, if $\gamma(1, \overrightarrow{v})$ is defined, then

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

This leads to the definition of the exponential map.

Definition 2.11.4 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

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

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

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 $\exp_p(\overrightarrow{v})$ is undefined for points on the circle of radius π.

Nevertheless, $\exp_p$ is always well-defined in a small open disk.

Lemma 2.11.5 *Given a surface $X\colon \Omega \to \mathbb{E}^3$, for every point $p = X(u,v)$ on X, there is some $\epsilon > 0$, 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 $\exp_p\colon B_\epsilon \to V$ is well defined and is a diffeomorphism.*

A neighborhood of p on X of the form $\exp_p(B_\epsilon)$ 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 coordinates 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 the 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 in 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.

Geodesic Lines, . . .
Covariant . . .
Home Page
Title Page

Lemma 2.11.6 *Given a surface* $X\colon \Omega \to \mathbb{E}^3$, *for every point* $p = X(u, v)$ *on* X, *there is some* $\epsilon > 0$ *and some open disk* B_ϵ *of center* $(0, 0)$ *such that for every* $q \in \exp_p(B_\epsilon)$, *for every geodesic* $\gamma\colon]-\eta, \eta[\to \mathbb{E}^3$ *in* $\exp_p(B_\epsilon)$ *such that* $\gamma(0) = p$ *and* $\gamma(t_1) = q$, *for every regular curve* $\alpha\colon [0, t_1] \to \mathbb{E}^3$ *on* X *such that* $\alpha(0) = p$ *and* $\alpha(t_1) = q$, *then*

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

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 2.11.6 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 2.11.7 *Given a surface $X\colon \Omega \to \mathbb{E}^3$, let $\alpha\colon I \to \mathbb{E}^3$ be a regular curve on X parameterized 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.

Nevertheless, we can't 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

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

and

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

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

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

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

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

does not change sign for $\eta \in] - \epsilon, \epsilon[$, 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 2.11.8 *Given a surface $X:\Omega \to \mathbb{E}^3$, 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*

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

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

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

Secondly, if we recall from lemma 2.7.1 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,

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

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, but this is beyond the scope of this course.

2.12. 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 thus, it is worth discussing anyway.

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 2.12.1 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, $\alpha\colon\,]-\epsilon, \epsilon[\, \to U$, 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

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

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 only depends 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 formulae (due to Gauss) for X_{uu}, X_{uv}, X_{vu}, and X_{vv}:

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

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

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

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

Definition 2.12.2 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 2.12.2 extends immediately to piecewise regular curves on a surface.

Definition 2.12.3 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 alternate definition of a geodesic.

Definition 2.12.4 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 2.12.5 *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.*

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 2.12.5, 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 w.r.t. the arc length $s = ct$, and we note that the parameter t of a geodesic is proportional to the arc length of α.

Lemma 2.12.6 *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, so that $w(t_0) = w_0$.*

Lemma 2.12.6 is an immediate consequence of standard results on ODE's. This lemma yields the notion of parallel transport.

Definition 2.12.7 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 2.12.6. 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, $\mathcal{X}(M)$, 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 $\mathcal{C}^\infty(M)$ of smooth functions $f\colon M \to \mathbb{R}$ on M.

Then, an *affine connection, D, on M* is a differentiable map,

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

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

(1) $D_{fX+gY} Z = f D_X Z + g D_Y Z$;

(2) $D_X(\lambda Y + \mu Z) = \lambda D_X Y + \mu D_X Z$;

(3) $D_X(fY) = f D_X Y + X(f) Y$,

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

Thus, an affine connection is $\mathcal{C}^\infty(M)$-linear in X, $\mathbb{R}$-linear in Y, and satisfies a "Leibnitz" type of law in Y.

For any chart $\varphi: U \to \mathbb{R}^m$, 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 2.12.8 *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 $\alpha: I \to M$ on M another vector field, DV/dt, along c, so 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 $Y \in \mathcal{X}(M)$, 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 Riemmanian 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, or any text on Riemannian geometry.