# Chapter 14

# Basics of The Differential Geometry of Surfaces

## 14.1. 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.

## 14.2. 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 $\Omega$ 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 $\Omega$.

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 $\dfrac{dC}{dt}(t)$ 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 $C\colon t \mapsto X(u(t), v(t))$ 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 14.2.1** A *surface patch* $X$, for short a *surface* $X$, is a map $X\colon \Omega \to \mathbb{E}^3$ where $\Omega$ 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* $\Omega$ 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 $\Omega$ is a *closed* subset of $\mathbb{R}^2$.

If $\Omega$ is a closed set, we assume that there is some open subset $U$ containing $\Omega$ 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

$$x = u(u^2 + v^2),$$
$$y = v(u^2 + v^2),$$
$$z = u^2v - v^3/3,$$

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,

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

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

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

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.

# 14.3. 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 quadratic form plays a major role in the theory of surfaces, and deserves an official definition.

**Definition 14.3.1** 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 $\varphi_I$ 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 $\varphi_I$ can be used to determine the angle of two curves passing through $p$, i.e., the angle $\theta$ 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 $\Omega$.

*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

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

This is the surface generated by a line parallel to the $xOy$ plane, touching the $z$ axis, and also touching 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 $0 \le v_1 \le 2\pi$ and $-2 \le u_1 \le 2$.

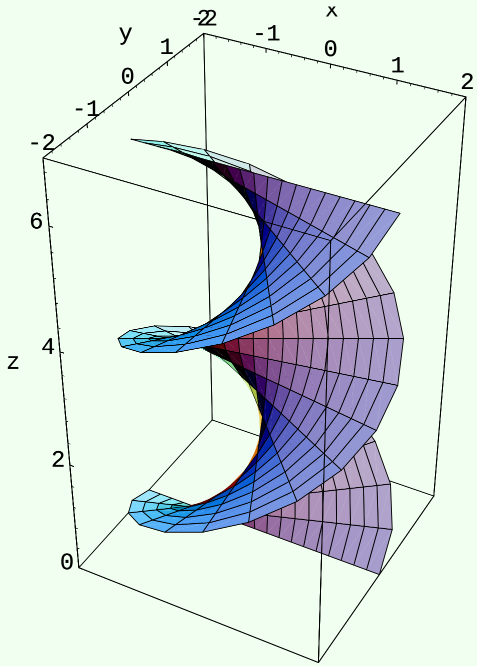


Figure 14.1: An helicoid

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

$$x = \cosh u_2 \cos v_2,$$
$$y = \cosh u_2 \sin v_2,$$
$$z = u_2.$$

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 $0 \le v_2 \le 2\pi$ and $-2 \le u_2 \le 2$.

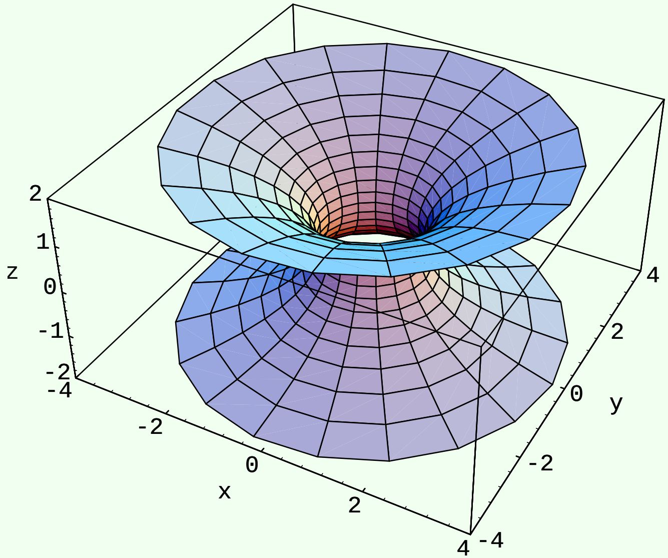


Figure 14.2: A catenoid

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 $2\pi$, 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.