Luigi Verdiani

Differential Geometry

Version : January 2017.

Notes for the students of Math 401/501 - University of Pennsylvania

Introduction

I will use these notes to add some support material for topics that are not covered in the textbook. This is a first version that will certainly contain mistakes....

Contents

Chapter 1

Curves

Isometries of the Euclidean space

We will mostly work in $\mathbb{R}^3$, but some of the definitions make sense in higher dimensions as well. Let

$$\mathbb{R}^n = \{(x_1, \ldots, x_n), x_i \in R, i = 1, \ldots, n\}$$

a point in $\mathbb{R}^n$ is just an ordered collection of real numbers. We know that, for $n = 3$, we can choose a coordinate system in the Euclidean space (i.e. an origin O, three axes and a unit to measure the distance) and a point in $x \in \mathbb{R}^3$ can be identified with a point P in the Euclidean space, or with a vector (O, P). In analogy with the tree dimensional case we will use the vector notation for points in $\mathbb{R}^n$. Moreover we can identify a point in $\mathbb{R}^n$ with a matrix:

$$(x_1, \ldots, x_n) \in \mathbb{R}^n \simeq \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M(n \times 1, \mathbb{R})$$

where $M(m \times n, \mathbb{R})$ is the set of matrices with m rows, n columns and real entries. Hence we have three different ways of looking at an element $x \in \mathbb{R}^n$.

In $\mathbb{R}^n$ we have a standard scalar (dot) product : for $x = (x_1, \ldots x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$

$$x \cdot y = x_1 y_1 + \ldots, x_n y_n = \sum_{i=1}^{n} x_i y_i.$$

The scalar product in $\mathbb{R}^n$ is a positive definite symmetric bilinear form i.e.

(i) $(x_1 + x_2) \cdot y = x_1 \cdot y + x_2 \cdot y, \qquad \forall x_1, x_2, y \in \mathbb{R}^n$
(ii) $(\lambda x) \cdot y = \lambda(x \cdot y), \qquad \forall x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$
(iii) $x \cdot (y_1 + y_2) = x \cdot y_1 + x \cdot y_2, \qquad \forall x, y_1, y_2 \in \mathbb{R}^n$
(iv) $x \cdot (\lambda y) = \lambda(x \cdot y), \qquad \forall x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$
(v) $x \cdot y = y \cdot x, \qquad \forall x, y \in \mathbb{R}^n$
(vi) $x \cdot x \geq 0, \qquad \forall x \in \mathbb{R}^n,$
(vii) $x \cdot x = 0 \iff x = 0$

Denote by $d(x, y)$ the Euclidean distance of x from y. For $n = 2$, $x \cdot x$ can be interpreted as the squared distance $d(x, 0)^2$ of x from the origin. In analogy with this case we define the Euclidean norm of x as

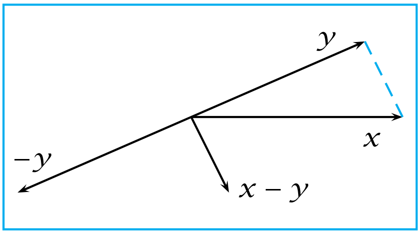

fig. 1.1

$$||x|| = \sqrt{x \cdot x}.$$

and $||x|| = d(x, 0)$ for every $x \in \mathbb{R}^n$. Moreover (see fig. 1.2) it is clear that

$$||x - y|| = d(x, y). \qquad (1.1)$$

And we can use the scalar product to recover informations about the distance of two points. But the scalar product gives also informations about the angles between vectors. For two nonzero vectors x and y we have

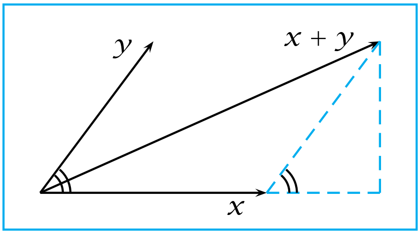

fig. 1.2

$$\begin{aligned} ||x + y||^2 &= (||x|| + ||y|| \cos(\widehat{xy}))^2 + \\ &+ (||y|| \sin(\widehat{xy}))^2 = \\ &= ||x||^2 + ||y||^2 + \\ &+ 2||x||\,||y|| \cos(\widehat{xy}). \end{aligned}$$

But, from the bilinearity of the scalar product we also have

$$||x + y||^2 = (x + y) \cdot (x + y) = ||x||^2 + ||y||^2 + 2x \cdot y$$

comparing the two expressions we find

$$x \cdot y = ||x||\,||y|| \cos(\widehat{xy}).$$

In particular the two vectors x and y are orthogonal if and only if $x \cdot y = 0$. If we replace y by $-y$, we have

$$d(x, y)^2 = ||x - y||^2 = ||x||^2 + ||y||^2 - 2\,x \cdot y.$$

If we identify x, y with $n \times 1$ matrices we also have

$$x \cdot y = x^t \cdot y$$

where x^t denotes the transpose of x and, on the right hand side $\cdot$ is the usual matrix product.

Corollary 1.1 *[Schwarz's inequality] For* $x, y \in \mathbb{R}^n$ *we have*

$$x \cdot y \leq ||x||\,||y||$$

and the equality holds if and only if x *and* y *are parallel;*

Definition 1.1 *A map* $f : \mathbb{R}^n \to \mathbb{R}^n$ *is an isometry if*

$$d(f(x), f(y)) = d(x, y)$$

for every $x, y \in \mathbb{R}^n$*.*

It follows that any triangle is mapped onto a congruent triangle, hence an isometry preserves the angles between vectors:

Lemma 1.1 *Let* $f : \mathbb{R}^n \to \mathbb{R}^n$ *be an isometry. Then*

$$\widehat{f(x)f(y)} = \widehat{xy}$$

for every $x, y \in \mathbb{R}^n$*.*

Moreover, by choosing $\delta = \epsilon$ in the definition of continuity:

Lemma 1.2 *Let* $f : \mathbb{R}^n \to \mathbb{R}^n$ *be an isometry. Then* f *is a continuous function.*

A first class of isometries is given by the translations: for a fixed $v \in \mathbb{R}^n$, we define a map $T_v : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T_v(x) = x + v.$$

Then

$$d(T_v(x), T_v(y)) = ||T_v(x) - T_v(y)|| = ||(x+v) - (y+v)|| = ||x-y|| = d(x, y)$$

hence T_v is an isometry. It is clear that T_v has an inverse, given by T_{-v}.

Lemma 1.3 *Let* $f, g : \mathbb{R}^n \to \mathbb{R}^n$ *be isometries. Then the composition* $g \circ f$ *is an isometry.*

Proof: For $x, y \in \mathbb{R}^n$ we have

$$d((g \circ f)(x), (g \circ f)(y)) = d(g(f(x)), g(f(y))) = d(f(x), f(y)) = d(x, y).$$

□

Now let $f : \mathbb{R}^n \to \mathbb{R}^n$ be an isometry and let $v = f(0)$. Then $\tilde{f} = T_{-v} \circ f$ is an isometry and

$$\tilde{f}(0) = (T_{-v} \circ f)(0) = T_{-v}(f(0)) = T_{-v}(v) = v - v = 0.$$

Hence $\tilde{f}$ fixes the origin and

$$T_v \circ \tilde{f} = T_v \circ (T_{-v} \circ f) = (T_{-v} \circ T_v) \circ f = f$$

and every isometry f is the composition of a translation and an isometry $\tilde{f}$ that fixes the origin. To understand the structure of the isometries of $\mathbb{R}^n$ we only have to study isometries that fix the origin.

Proposition 1.1 *A map $f : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry that fixes the origin if and only if f preserves the scalar product.*

Proof: Let f be an isometry such that $f(0) = 0$. Then, for $x \in \mathbb{R}^n$:

$$||x|| = d(x, 0) = d(f(x), f(0)) = d(f(x), 0) = ||f(x)||.$$

Using the fact that the angles are preserved by f we have, for $x, y \in \mathbb{R}^n$

$$f(x) \cdot f(y) = ||f(x)||\,||f(y)|| \cos(\widehat{f(x)f(y)}) = ||x||\,||y|| \cos(\widehat{xy}) = x \cdot y.$$

hence f preserves the scalar product. Conversely, assuming that f preserves the scalar product

$$||x||^2 = x \cdot x = f(x) \cdot f(x) = ||f(x)||^2$$

and f preserves the norm of any vector $x \in \mathbb{R}^n$. In particular $||f(0)|| = ||0|| = 0$ hence $f(0) = 0$. Finally

$$\begin{aligned} d(f(x), f(y))^2 &= ||f(x) - f(y)||^2 = \\ &= ||f(x)||^2 + ||f(y)||^2 - 2f(x) \cdot f(y) = \\ &= ||x||^2 + ||y||^2 - 2x \cdot y = ||x - y||^2 = d(x, y)^2 \end{aligned}$$

and f is an isometry. □

If $A \in M(n \times n, \mathbb{R})$ we can define a linear map $f_A : \mathbb{R}^n \to \mathbb{R}^n$ by

$$f_A(x) = A \cdot x$$

and we consider a special class of maps, defined by orthogonal matrices

$$A \in O(n) = \{A \in M(n \times n, \mathbb{R}) : A^t \cdot A = I\}.$$

We have

$$f_A(x) \cdot f_A(y) = (A \cdot x)^t \cdot (A \cdot y) = x^t \cdot A^t \cdot A \cdot y = x \cdot y$$

hence, when $A \in O(n)$, the map f_A preserves the scalar product. Since $f_A(0) = 0$ because f_A is linear, we have, form the previous Proposition, that f_A is an isometry.

Proposition 1.2 *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be an isometry such that $f(0) = 0$ then f is a linear map and $f = f_A$ for some orthogonal matrix $A \in O(n)$.*

Proof: We first prove that f is linear. Let $x, y \in \mathbb{R}^n$. As we observed before, the image of a triangle with vertices $0, x, x + y$ is a triangle congruent to the original one. The same is true for the triangle with vertices $0, y, x + y$. It follows easily that the parallelogram with vertices $0, x, x + y, y$ is mapped onto a congruent parallelogram with vertices $0, f(x), f(x + y), f(y)$. In particular $f(x + y)$ is the vector sum of $f(x)$ and $f(y)$, i.e.

$$f(x + y) = f(x) + f(y) \tag{1.2}$$

for every $x, y \in \mathbb{R}^n$. Thus we proved that f satisfies the first part of the definition of a linear map. Now we want to prove that

$$f(\lambda x) = \lambda f(x) \tag{1.3}$$

for every $\lambda \in R$ and $x \in \mathbb{R}^n$. From (1.2) we get

$$f(2x) = f(x + x) = f(x) + f(x) = 2f(x).$$

By induction on n it follows that

$$f(nx) = nf(x)$$

for every $n \in \mathbb{N}$. From (1.2) and the fact that $f(0) = 0$ it follows that $f(-x) = -f(x)$. For $n \in N$ we have then

$$f((-n)x) = f(n(-x)) = nf(-x) = -nf(x).$$

hence (1.3) is satisfied for $\lambda \in \mathbb{Z}$. Let $p, q \in \mathbb{Z}$ with $q \neq 0$. Then

$$\frac{p}{q} f(x) = \frac{p}{q} f(q\frac{x}{q}) = pf(\frac{x}{q}) = f(\frac{p}{q}x)$$

and (1.3) holds for $\lambda \in \mathbb{Q}$. If $\lambda \in \mathbb{R}$ we choose a sequence $\lambda_n \in \mathbb{Q}$ that converges to λ . Then, using the continuity of f

$$\lambda f(x) = \lim_{n\to\infty} \lambda_n f(x) = \lim_{n\to\infty} f(\lambda_n x) = f(\lim_{n\to\infty} \lambda_n x) = f(\lambda x)$$

and f is a linear map, we can write $f = f_A$ for some matrix $A \in M(n \times n, \mathbb{R})$. Since f is an isometry such that $f(0) = 0$ we have

$$x \cdot y = f_A(x) \cdot f_A(y) = (A \cdot x)^t \cdot (A \cdot y) = x^t \cdot A^t \cdot A \cdot y$$

for every $x, y \in \mathbb{R}^n$. Denote by e_i the standard basis of $\mathbb{R}^n$ and by δ_{ij} the Kronecker symbol. Then $e_i \cdot e_j = \delta_{ij}$. Let $B = A^t \cdot A$ then it is easy to show that $e_i^t \cdot B \cdot e_j = b_{ij}$. Hence $A^t \cdot A = B = I$ and $A \in O(n)$. □

Let $A \in O(n)$. Then we have

$$1 = \det(I) = \det(A^t \cdot A) = \det(A^t)\det(A) = \det(A)^2$$

hence $\det(A) = \pm 1$. If $\det(A) = 1$ we say that the isometry f_A is orientation preserving , otherwise f_A is orientation reversing.

Let $x(t) = (x_1(t), \dots, x_n(t))$, $y(t) = (y_1(t), \dots, y_n(t)) \in \mathbb{R}^n$ be two families of vectors depending on a real parameter t. Assume that the vectors are of class C^1 i.e. the functions $x_i(t), y_i(t)$ are differentiable for $i = 1, \dots, n$. Then the scalar product $x(t) \cdot y(t)$ is a real valued function and we may differentiate it:

$$\begin{aligned} (x(t) \cdot y(t))' &= \left(\sum_{i=1}^{n} x_i(t) y_i(t) \right)' = \sum_{i=1}^{n} x_i(t)' y_i(t) + \sum_{i=1}^{n} x_i(t) y_i(t)' = \\ &= x(t)' \cdot y(t) + x(t) \cdot y(t)' \end{aligned} \tag{1.4}$$

Remark 1.1 *As an immediate consequence we have:*

(i) *If $||x(t)||$ is constant then $x(t) \cdot x'(t) = 0$.*

(ii) *If $x(t)$ and $y(t)$ are orthogonal for all t then $x'(t) \cdot y(t) = -x(t) \cdot y'(t)$.*

Curves

We have two different points of view for curves: the first one is more 'geometric', we are interested in a geometric locus given by is a one dimensional subset of $\mathbb{R}^n$. The second one is more related to physics: a curve is the trajectory of a moving point particle, meaning that we keep track of the way the particle moves (for instance the speed).

Example 1.1 *A definition of the unit circle in $\mathbb{R}^2$ that fits into the first scheme is*

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

while

$$\alpha_1(t) = (\cos(t), \sin(t)), \qquad t \in [0, 2\pi]$$

$$\alpha_2(t) = (\cos(2t), \sin(2t)), \qquad t \in [0, \pi] \tag{1.5}$$

define the same trajectory but the speed of the particle is higher in α_2 than in α_1.

At the end we are interested in the 'shape' of a curve but the second point of view will give us more informations and tools useful to study it.

Definition 1.2 *A map $f : \mathbb{R}^m \to \mathbb{R}^n$ is of class C^k if*

$$f(x_1, \dots, x_m) = (f_1(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m))$$

and each of the first k derivatives of the functions $f_i : \mathbb{R}^m \to \mathbb{R}$, $i = 1, \dots, n$, is continuous.

Definition 1.3 *A curve of class C^k in $\mathbb{R}^n$ is a map $\alpha : I \to \mathbb{R}^n$ of class C^k where $I \subset \mathbb{R}$ is an interval (possibly unbounded). The trace of α is the image $\alpha(I)$ of the interval I.*

If there exists a plane $\pi \subset \mathbb{R}^n$ such $\alpha(I) \subset \pi$ then we will say that α is a plane curve. We can always move a plane by an isometry and assume that the plane passes through the origin and rotate it so that it coincides with the set $\{(x_1, \dots, x_n) \in \mathbb{R}^n : x_3 = \dots, x_n = 0\}$. In most cases we will identify this plane with $\mathbb{R}^2$.

If α is of class C^k then, for $1 \le i \le k$, the i-th derivative of α is

$$\alpha^{(i)} = (\alpha_1^{(i)}, \dots, \alpha_n^{(i)}).$$

Note that

$$\begin{aligned}
\alpha'(t_0) &= (\alpha_1'(t_0), \dots, \alpha_n'(t_0)) = \\
&= \left(\lim_{t \to t_0} \frac{\alpha_1(t) - \alpha_1(t_0)}{t - t_0}, \dots, \lim_{t \to t_0} \frac{\alpha_n(t) - \alpha_n(t_0)}{t - t_0}\right) = \\
&= \lim_{t \to t_0} \left(\frac{\alpha_1(t) - \alpha_1(t_0)}{t - t_0}, \dots, \frac{\alpha_n(t) - \alpha_n(t_0)}{t - t_0}\right) = \\
&= \lim_{t \to t_0} \frac{1}{t - t_0} (\alpha_1(t) - \alpha_1(t_0), \dots, \alpha_n(t) - \alpha_n(t_0)) = \\
&= \lim_{t \to t_0} \frac{\alpha(t) - \alpha(t_0)}{t - t_0}
\end{aligned}$$

for every t, $\alpha(t) - \alpha(t_0)$ is a vector, and we get a vector in the limit, called the tangent vector to α at $t = t_0$. The norm of $\alpha'(t)$ is the speed of α. Similarly $\alpha''(t)$ is the acceleration vector of α.

Definition 1.4 *A curve $\alpha : I \to \mathbb{R}^n$ of class C^k is k-regular if $\alpha'(t), \dots, \alpha^{(k)}(t)$ are linearly independent for every $t \in I$.*

If α is 1-regular then we will just say that α is regular. A curve α is then regular if and only if $\alpha'(t) \ne 0$ for every $t \in I$.

Example 1.2 *The cuspidal cubic is a curve $\alpha : \mathbb{R} \to \mathbb{R}^2$ defined by*

$$\alpha(t) = (t^2, t^3)$$

The trace of this curve (see fig. 1.3) can also be defined as

$$\{(x, y) \in \mathbb{R}^2 : y^2 - x^3 = 0\}$$

in fact a cubic is a curve defined as the zero locus of a polynomial of degree three. This cubic is not regular since $\alpha'(t) = (2t, 3t^2)$ vanishes at $t = 0$. This is the only singular point of α and there is no tangent vector for $t = 0$. The curve α is smooth, but, in a neighborhood of $t = 0$, the trace of α is not the graph of a smooth function (in this case a function of the y variable), we have $x = y^{\frac{3}{2}}$.

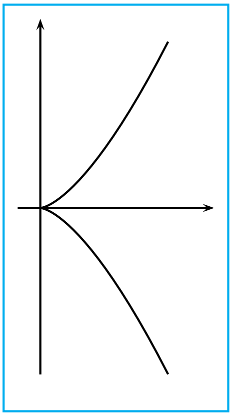

fig. 1.3

Example 1.3 *The nodal cubic is a curve $\alpha : \mathbb{R} \to \mathbb{R}^2$ defined by*

$$\alpha(t) = (t^2 - 1, t(t^2 - 1))$$

The trace of this curve (see fig. 1.4) can also be defined as

$$\{(x, y) \in \mathbb{R}^2 : y^2 - x^3 - x^2 = 0\}$$

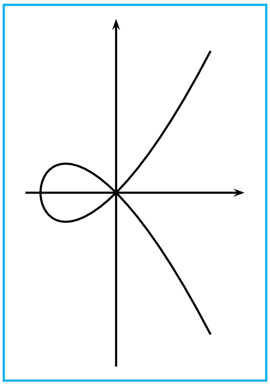

fig. 1.4

The image of the map α is not injective, we have $\alpha(1) = \alpha(-1) = (0, 0)$ hence it does not make sense to speak about the tangent vector at a point of the trace of a curve. In this example it is possible to describe the zero set of a polynomial by intersecting with a straight line: let $y = tx$ be a line through the origin, the intersection with $y^2 - x^3 - x^2 = 0$ occurs when $x^2(-x + t^2 - 1) = 0$ i.e. $x = t^2 - 1$ and $y = t(t^2 - 1)$. This method will not work in general since a straight line through the origin will intersect the locus in more than one point.

Example 1.4 *The Folium of Decartes is a curve $\alpha : \mathbb{R} \to \mathbb{R}^2$ defined by*

$$\alpha(t) = \left(\frac{3t}{1 + t^3}, \frac{3t^2}{1 + t^3} \right).$$

The curve α is injective, but the trace of this curve (see fig. 1.5) is not locally homeomorphic to an interval in the real line. Every subset X of $\mathbb{R}^n$ has a topology induced by the topology of $\mathbb{R}^n$. The open sets in X are of the form $X \cap U$ where U is open in $\mathbb{R}^n$. As t grows, the points of the curve accumulate close to the origin and every (small enough) neighborhood of the origin in the trace of α is disconnected.

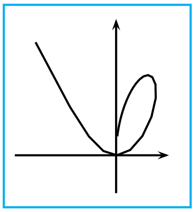

fig. 1.5

Now we move to more standard examples :

Example 1.5 *Let $p, q \in \mathbb{R}^n$. The straight line through p and q is defined by*

$$\alpha(t) = q + (1-t)(p-q).$$

This is a regular curve and $\alpha'(t)$ is constant equal to $q - p$.

Example 1.6 *Let a, b be positive real numbers and let $p = (x_0, y_0) \in \mathbb{R}^2$. Then*

$$\alpha(t) = (x_0 + a\cos(t), y_0 + b\sin(t))$$

is an ellipse with center in p. If $a = b$ we have a circle.

Example 1.7 *A cycloid is the path traced by a point p on the boundary of a circle (say of radius 1) as the circle rolls (without slipping) along a straight line. Let t be the angle of the circle's rotation and assume that, for $t = 0$, the circle is centered in $(0,1)$ and the point p is $(0,0)$. Then the coordinates of the point p (fig. 1.6) at time t are given by*

$$\alpha(t) = (t - \sin(t), 1 - \cos(t)).$$

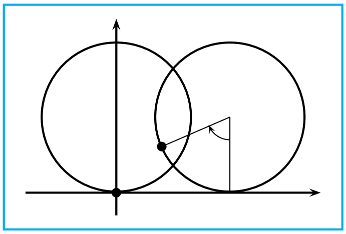

fig. 1.6

The resulting curve is not regular, we have $\alpha'(t) = (1 - \cos(t), \sin(t))$ and this vector vanishes for $t = 2k\pi$, $k \in \mathbb{Z}$, i.e. every time the curve intersects the x axis. The cycloid was studied, in particular, during the XVII century when it was discovered that it has remarkable properties, try to google the words tautochrone or brachystocrone...

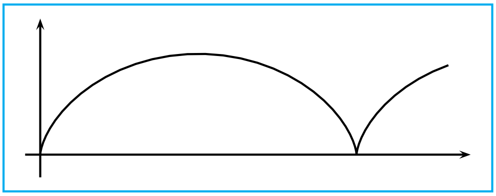

fig. 1.7

Example 1.8 *An Astroid is a curve traced out by a point p on the circumference of one circle (of radius a) as that circle rolls without slipping on the inside of a second circle having four times the radius of the first. Let t be the angle that measures the rotation of the center of the small circle and denote by θ the angle that measures the rotation of the point p on the small circle. Since the small circle rotates without slipping we have $at = \frac{a}{4}\theta$. Hence $\theta = 4t$. Assume that, for $t = 0$, the small circle is centered in $(\frac{3}{4}a, 0)$ and the point p is $(a, 0)$.*

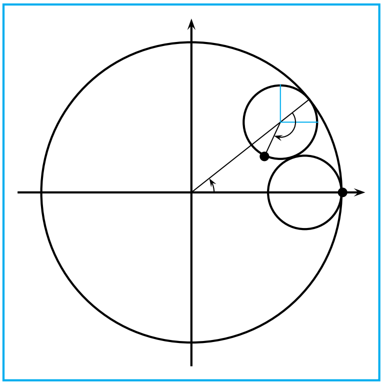

fig. 1.8

We describe the trajectory of p as sum of two vectors. The first one describe the trajectory of the center of the small circle. The second one the coordinates of the point p (fig. 1.8) at time t, with respect to a frame centered at the center of the small circle, these are given by

$$\left(\frac{a}{4}\cos(3t), -\frac{a}{4}\sin(3t)\right)$$

since the coordinates of the center of the small circle are

$$\left(\frac{3a}{4}\cos(t), \frac{3a}{4}\sin(t)\right)$$

the coordinates of p are given by the sum of these two vectors and parametric equation of the astroid is

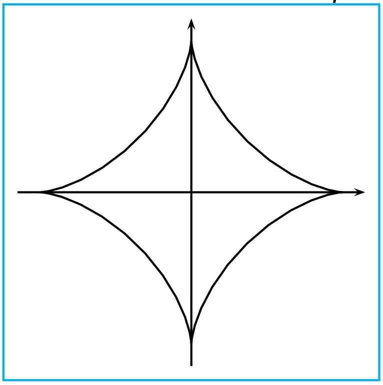

fig. 1.9

$$\alpha(t) = \left(\frac{3a}{4}\cos(t) + \frac{a}{4}\cos(3t), \frac{3a}{4}\sin(t) - \frac{a}{4}\sin(3t)\right)$$

Using trigonometric formulas it is possible to show that

$$\alpha(t) = (a\cos(t)^3, a\sin(t)^3).$$

The resulting curve is not regular, we have

$$\alpha'(t) = (-3a\cos(t)^2\sin(t), 3a\sin(t)^2\ \cos(t))$$

and this vector vanishes for $t = k\frac{\pi}{2}$, $k \in \mathbb{Z}$.

Arclength

Definition 1.5 *Two curves $\alpha : I \to \mathbb{R}^n$, $\beta : J \to \mathbb{R}^n$ differ be the parametrization if there exists a diffeomorphism $\phi : J \to I$ such that $\beta(s) = \alpha(\phi(s))$. We will also say that β is a reparametrization of α.*

In general, if α is not smooth, it is enough to require that ϕ has the same regularity of α (it follows from the inverse function theorem that ϕ^{-1} has the same regularity as well). In particular two curves that differ by the parametrization have the same trace.

Note that $I = [a, b]$ and $J = [c, d]$ are intervals in the real line. A diffeomorphism $\phi : J \to I$ must be monotone and maps the boundary of the interval J into the boundary of the interval I. Hence we have two possible cases: $\phi(c) = a$ and $\phi(d) = b$ (if ϕ is increasing) or $\phi(c) = b$ and $\phi(d) = a$ (if ϕ is decreasing). Since we are mainly interested in the geometric properties of the trace of a curve α , we will look for the best possible parametrization of α.

Definition 1.6 *Let $\alpha : [a,b] \to \mathbb{R}^n$ be a curve of class C^1. Then we define the length of α by*

$$L(\alpha) = \int_a^b ||\alpha'(t)||\, dt.$$

The length of a curve does not depend on the parametrization:

‖ **Proposition 1.3** *Let β be a reparametrization of a curve α. Then $L(\alpha) = L(\beta)$.*

Proof: Let $\alpha : [a,b] \to \mathbb{R}^n$ and $\beta : [c,d] \to \mathbb{R}^n$ and let $\phi : [c,d] \to [a,b]$ be a diffeomorphism. Then ϕ is monotone and we have two possible cases:

(i) ϕ is increasing:

$$\begin{aligned} L(\beta) &= \int_c^d ||\beta'(s)||\, ds = \int_c^d ||\alpha(\phi(s))'||\, ds = \\ &= \int_c^d ||\alpha'(\phi(s))\phi'(s)||\, ds = \int_c^d ||\alpha'(\phi(s))|||\phi'(s)|\, ds = \\ &= \int_c^d ||\alpha'(\phi(s))||\phi'(s)\, ds = \int_a^b ||\alpha'(t)||\, dt = L(\alpha) \end{aligned}$$

(ii) ϕ is decreasing:

$$\begin{aligned} L(\beta) &= \int_c^d ||\beta'(s)||\, ds = \int_c^d ||\alpha(\phi(s))'||\, ds = \\ &= \int_c^d ||\alpha'(\phi(s))\phi'(s)||\, ds = \int_c^d ||\alpha'(\phi(s))|||\phi'(s)|\, ds = \\ &= -\int_c^d ||\alpha'(\phi(s))||\phi'(s)\, ds = -\int_b^a ||\alpha'(t)||\, dt = \\ &= \int_a^b ||\alpha'(t)||\, dt = L(\alpha) \end{aligned}$$

□

Let $U \subset \mathbb{R}^n$ be a connected subset of $\mathbb{R}^n$. Then we define the distance between two points $p, q \in U$ by:

$$d(p,q) = \inf\{L(\alpha) : \alpha : [a,b] \to U, \alpha(a) = p, \alpha(b) = q\}.$$

Sometimes it is not possible to find a minimizing curve: let $U = \mathbb{R}^2 \setminus \{(0,0\}$ and $p = (-1,0)$, $q = (0,1)$. The minimizing curve between p and q in $\mathbb{R}^2$ is the segment joining the two points, but the trace of this curve does not lie in U, hence we can only approximate it with curves in U and get it as a limit. In the case of $\mathbb{R}^n$ this definition is consistent with the one given in a previous section:

Proposition 1.4 *Let $p, q \in \mathbb{R}^n$. Then the distance between p, q is the length of the line segment joining p and q.*

Proof: Let $\alpha : [0,1] \to \mathbb{R}^n$ be given by $\alpha(t) = q + (t-1)(q-p)$ then $\alpha(0) = p$ and $\alpha(1) = q$ and α is a parametrization of the line segment. We have

$$L(\alpha) = \int_0^1 ||q-p||\, dt = ||q-p|| \int_0^1 dt = ||q-p||.$$

Let $\beta : [0,1] :\to \mathbb{R}^n$ be any curve such that $\beta(0) = p$ and $\beta(1) = q$. Let $v = (q-p)/||q-p||$ (hence v is constant unit vector). We have

$$\int_0^1 \beta'(t)\cdot v\, dt = \int_0^1 (\beta(t)\cdot v)'\, dt = \beta(1)\cdot v - \beta(0)\cdot v = (q-p)\cdot v = ||q-p|| = L(\alpha).$$

and, using the Schwarz inequality

$$\int_0^1 \beta'(t)\cdot v\, dt \le \int_0^1 ||\beta'(t)||\,||v||\, dt = \int_0^1 ||\beta'(t)||\, dt = L(\beta).$$

Hence $L(\beta) \ge L(\alpha)$ for any other curve β and α is a minimizer according to the new definition of distance. □

Definition 1.7 *Let $\alpha : I \to \mathbb{R}^n$ be a C^1 curve and let $t_0 \in I$. The arclength function relative to t_0 is defined by:*

$$s(t) = \int_{t_0}^t ||\alpha'(t)||\, dt.$$

In other words the arclength function measures (up to sign) the length of the curve segment between $\alpha(t_0)$ and $\alpha(t)$.

It is clear the the arclength function is differentiable and we have

$$s'(t) = ||\alpha'(t)|| \ge 0$$

if we assume that α is a regular curve then $s(t)$ is a differentiable strictly increasing function and we have a differentiable inverse $t(s) : J \to I$ for some interval $J \subset \mathbb{R}$. It is then possible to consider the new parametrization of α given by $\beta : J \to \mathbb{R}^n$:

$$\beta(s) = \alpha(t(s)).$$

The chain rule for the derivative of a function then shows that

$$\beta'(s) = \alpha'(t(s))t'(s) = \frac{1}{s'(t)}\alpha'(t(s)) = \frac{1}{||\alpha'(t(s))||}\alpha'(t(s))$$

so that $||\beta'(s)|| = 1$. In other words:

Proposition 1.5 *If $\alpha : I \to \mathbb{R}^n$ is a regular curve of class C^1, then it is possible to find a parametrization β of α such that $||\beta'|| = 1$. We say that α is then parametrized by arclength.*

The parametrization by arclength has the property that the length of the curve segment between two points $\beta(s)$ and $\beta(s_0)$, with $s > s_0$ is $s - s_0$. If $I = [a, b]$ we will often assume that $t_0 = a$ so that $J = [0, L]$ where L is the length of the curve α.

Example 1.9 *Consider the circle $\alpha(t) = (x_0 + r\cos(t), y_0 + r\sin(t))$ for $t \in [0, 2\pi]$. Then $\alpha'(t) = (-r\sin(t), r\cos(t))$ and $||\alpha'(t)|| = r$. Hence*

$$s(t) = \int_0^L r\,dt = r\,t$$

so that $t(s) = \frac{s}{r}$ and $\beta(s) = (x_0 + r\cos(\frac{s}{r}), y_0 + r\sin(\frac{s}{r}))$ for $s \in [0, 2\pi]$.

Example 1.10 *Consider the cycloid $\alpha(t) = (t - \sin(t), 1 - \cos(t))$. Then $\alpha'(t) = (1 - \cos(t), \sin(t))$. If $t \in [0, 2\pi]$ then α is regular in the interior of the interval and we have*

$$\begin{aligned} s(t) &= \int_0^t \sqrt{2 - 2\cos(t)}\,dt = \int_0^t 2\sqrt{\frac{1-\cos(t)}{2}}\,dt = \\ &= \int_0^t 2\sin(\frac{t}{2})\,dt = 4(1 - \cos(\frac{t}{2})). \end{aligned}$$

The length of the curve, in the interval $[0, 2\pi]$ is then equal to 8, and $t(s) = 2\arccos(1 - \frac{s}{4})$ for $s \in [0, 8]$.

In general it is difficult to compute the arclength function for a given function, this requires finding a primitive and an inverse. The importance of the arclength is mainly theoretical. In some applications it is possible to use a numerical approximation.

The Frenet frame

From now on we will only consider curves in $\mathbb{R}^3$. There is an n-dimensional analogue for all the statements in this section, and the proofs are substantially the same.

Let $\alpha : I \to \mathbb{R}^3$ be a 2-regular (or bi-regular) curve. We normalize the tangent vector $\alpha'(t)$ in order to get a unit vector (this is automatic if α is parametrized by arclength, bur here we are not assuming that). We let

$$\underline{t} = \frac{1}{||\alpha'(t)||}\alpha'(t)$$

then

$$\begin{aligned}
\frac{d\underline{t}}{dt} &= \frac{d}{dt}\left(\frac{1}{\sqrt{\alpha'(t)\cdot\alpha'(t)}}\alpha'(t)\right) = \\
&= \frac{1}{||\alpha'(t)||}\alpha''(t) - \frac{1}{2}\frac{2(\alpha'(t)\cdot\alpha''(t))}{(\alpha'\cdot\alpha')^{32}}\alpha'(t) = \\
&= \frac{1}{||\alpha'(t)||}\alpha''(t) - \frac{\alpha'(t)\cdot\alpha''(t)}{||\alpha'(t)||^3}\alpha'(t) = \\
&= \frac{1}{||\alpha'(t)||}\left(\alpha''(t) - \frac{\alpha'(t)\cdot\alpha''(t)}{||\alpha'(t)||^2}\alpha'(t)\right) \qquad (1.6)
\end{aligned}$$

and we define the normal vector by

$$\underline{n} = \frac{1}{||\frac{d\underline{t}}{dt}||}\frac{d\underline{t}}{dt}|.$$

Then $\underline{n}$ is a unit vector (here we are using the fact that α is bi-regular) and, since $\underline{t}$ has constant length, it follows from Remark 1.1 that $\underline{t}$ and $\underline{n}$ are orthogonal. We define the bi-normal vector by

$$\underline{b} = \underline{t} \times \underline{n}$$

where $\times$ denotes the cross product. Since $\underline{t}$ and $\underline{n}$ are unit and orthogonal we have that $\underline{b}$ is a unit vector orthogonal to the plane spanned by $\underline{t}$ and $\underline{b}$

Definition 1.8 *Let $\alpha : I \to \mathbb{R}^3$ be a bi-regular curve. The orthonormal frame $\{\underline{t}, \underline{n}, \underline{b}\}$ is the Frenet frame of $\alpha(t)$.*

The Frenet frame is a family of positive orthonormal basis of $\mathbb{R}^3$ parametrized by the points of α. Let $\omega_{12} = ||\frac{d\underline{t}}{dt}||$. Then, by definition

$$\frac{d\underline{t}}{dt} = \omega_{12}\,\underline{n}$$

Since $\underline{b}$ has constant length we have $\frac{d\underline{b}}{dt} \perp \underline{b}$. Since $\underline{b}\cdot\underline{t} = 0$ it follows from Remark 1.1 that

$$\frac{d\underline{b}}{dt}\cdot\underline{t} = -\frac{d\underline{t}}{dt}\cdot\underline{b} = 0$$

hence $e\frac{d\underline{b}}{dt}$ is parallel to $\underline{n}$ and we let

$$\frac{d\underline{b}}{dt} = \omega_{32}\,\underline{n}$$

since $\underline{t}\cdot\underline{n} = \underline{b}\cdot\underline{n} = 0$ we obtain similarly

$$\frac{d\underline{n}}{dt}\cdot\underline{t} = -\omega_{12}, \qquad \frac{d\underline{n}}{dt}\cdot\underline{b} = -\omega_{32}.$$

Summing up we obtained the Frenet's equations

$$\begin{cases} \frac{d\underline{t}}{dt} = \omega_{12}\,\underline{n} \\ \frac{d\underline{n}}{dt} = -\omega_{12}\,\underline{t} - \omega_{32}\,\underline{b} \\ \frac{d\underline{b}}{dt} = \omega_{32}\,\underline{n} \end{cases}$$

or, in matrix form

$$\begin{pmatrix} \frac{d\underline{t}}{dt} \\ \frac{d\underline{n}}{dt} \\ \frac{d\underline{b}}{dt} \end{pmatrix} = \begin{pmatrix} 0 & \omega_{12} & 0 \\ -\omega_{12} & 0 & -\omega_{32} \\ 0 & \omega_{32} & 0 \end{pmatrix} \cdot \begin{pmatrix} \underline{t} \\ \underline{n} \\ \underline{b} \end{pmatrix}.$$

The matrix

$$\Omega = \begin{pmatrix} 0 & \omega_{12} & 0 \\ -\omega_{12} & 0 & -\omega_{32} \\ 0 & \omega_{32} & 0 \end{pmatrix}$$

is skew-symmetric and this fact will be important later. The matrix Ω depends on the parametrization of $\alpha(t)$ but if we let

$$\mathrm{k}(t) = \frac{\omega_{12}}{||\alpha'(t)||}, \qquad \tau(t) = \frac{\omega_{32}}{||\alpha'(t)||}$$

the functions $\mathrm{k}(t)$, the curvature of α and $\tau(t)$, the torsion of α, then these functions are independent on the parametrization (we will prove it later). Note that, since $\omega_{12} > 0$, the curvature $\mathrm{k}(t)$ is always positive. From (1.6) we derive

$$\begin{aligned} \alpha''(t) &= ||\alpha'(t)||\,\frac{d\underline{t}}{dt} + \frac{\alpha'(t)\cdot\alpha''(t)}{||\alpha'(t)||}\,\frac{1}{||\alpha'(t)||}\alpha'(t) = \\ &= ||\alpha'(t)||^2\ \mathrm{k}(t)\,\underline{n} + \frac{\alpha'(t)\cdot\alpha''(t)}{||\alpha'(t)||}\,\underline{t} = \\ &= ||\alpha'(t)||^2\ \mathrm{k}(t)\underline{n} + ||\alpha'(t)||'\,\underline{t}. \end{aligned} \tag{1.7}$$

i.e. the decomposition of the acceleration of α into tangent and normal component. if the curve is parametrized by arclength then wee see that the curvature measures the variation of the velocity vector.

The curvature and the torsion of α will be the key quantities needed to understand the structure of α. We will now derive more efficient formulas to compute them.

Lemma 1.4 *Let $x, y \in \mathbb{R}^n$ be nonzero vectors. Then*

$$||x \times y||^2 = ||x||^2\,||y||^2 - (x \cdot y)^2\,.$$

Proof:

$$
\begin{aligned}
||x \times y||^2 &= ||x||^2\,||y||^2 \sin(\widehat{xy})^2 = ||x||^2\,||y||^2\,(1-\cos(\widehat{xy})^2) = \\
&= ||x||^2\,||y||^2\,(1-\left(\frac{x\cdot y}{||x||\,||y||}\right)^2) = ||x||^2\,||y||^2 - \left(x\cdot y\right)^2.
\end{aligned}
$$

□

Proposition 1.6 *Let $\alpha : I \to \mathbb{R}$ be a bi-regular curve. Then the curvature of $\alpha(t)$ is given by the following formula.*

$$
\mathrm{k}(t) = \frac{||\alpha'(t) \times \alpha''(t)||}{||\alpha'(t)||^3}.
$$

Proof: We have, using (1.6)

$$
\begin{aligned}
||\frac{d\underline{t}}{dt}||^2 &= \frac{d\underline{t}}{dt}\cdot\frac{d\underline{t}}{dt} = \frac{1}{||\alpha'(t)||^2}\left(\alpha''(t) - \frac{\alpha'(t)\cdot\alpha''(t)}{||\alpha'(t)||^2}\right)\cdot\left(\alpha''(t) - \frac{\alpha'(t)\cdot\alpha''(t)}{||\alpha'(t)||^2}\right) = \\
&= \frac{1}{||\alpha'(t)||^2}\left(\alpha''(t)\cdot\alpha''(t) - 2\frac{(\alpha'(t)\cdot\alpha''(t)^2}{||\alpha'(t)||^2} + \frac{(\alpha'(t)\cdot\alpha''(t)^2}{||\alpha'(t)||^2}\right) = \\
&= \frac{1}{||\alpha'(t)||^2}\left(||\alpha''(t)||^2 - \frac{(\alpha'(t)\cdot\alpha''(t)^2}{||\alpha'(t)||^2}\right) = \\
&= \frac{1}{||\alpha'(t)||^4}\left(||\alpha''(t)||^2\,||\alpha'(t)||^2 - (\alpha'(t)\cdot\alpha''(t)^2\right) = \\
&= \frac{1}{||\alpha'(t)||^4}||\alpha'(t)\times\alpha''(t)||^2
\end{aligned}
$$

Hence

$$
||\frac{d\underline{t}}{dt}|| = \frac{1}{||\alpha'(t)||^2}||\alpha'(t)\times\alpha''(t)|| \tag{1.8}
$$

and we can conclude using the definition of ω_{12}. □

Note that this formula makes sense also if the curve is only 1-regular.

Corollary 1.2 *Let $\alpha : I \to \mathbb{R}$ be a bi-regular curve. Then the curvature of $\alpha(t)$ does not depend on the parametrization of α.*

Proof: Let $\beta(s) = \alpha(\phi(s))$ be a different parametrization of α. Then

$$
\beta'(s) = \alpha'(\phi(s))\,\phi'(s), \qquad \beta''(s) = \alpha''(\phi(s))\,\phi'(s)^2 + \alpha'(\phi(s))\,\phi'(s).
$$

Then

$$\begin{aligned}
\mathrm{k}(s) &= \frac{||\beta'(s)\times\beta''(s)||}{||\beta'(s)||^3} = \frac{||\phi'(s)^3\ (\alpha'(\phi(s))\times\alpha''(\phi(s)))\,||}{||\alpha'(\phi(s))||^3\phi'(s)^3} = \\
&= \frac{||\alpha'(\phi(s))\times\alpha''(\phi(s))||}{||\alpha'(\phi(s))||^3}.
\end{aligned}$$

□

Now we want to derive a similar expression for the torsion of α:

Proposition 1.7 *Let $\alpha : I \to \mathbb{R}$ be a bi-regular curve. Then the torsion of $\alpha(t)$ is given by the following formula.*

$$\tau(t) = \frac{\alpha'\times\alpha'''\cdot\alpha''}{\|\alpha'(t)\times\alpha''(t)\|^2}.$$

Proof: Using (1.7) we have

$$\alpha'(t)\times\alpha''(t) = ||\alpha'(t)||\,\underline{t}\times\left(||\alpha'(t)||^2\ \mathrm{k}(t)\underline{n} + ||\alpha'(t)||'\,\underline{t}\right) = ||\alpha'(t)||^3\ \mathrm{k}(t)\,\underline{b}.$$

And

$$\underline{b} = \frac{1}{||\alpha'(t)\times\alpha''(t)||}\,\alpha'(t)\times\alpha''(t).$$

Then, using the fact that the cross product of two parallel vectors vanishes,

$$\frac{d\underline{b}}{dt} = \frac{1}{||\alpha'(t)\times\alpha''(t)||}\,\alpha'(t)\times\alpha'''(t) + \left(\frac{1}{||\alpha'(t)\times\alpha''(t)||}\right)'\,\alpha'(t)\times\alpha''(t).$$

And, since the triple product of three vectors that lie on the same plane is zero (and using (1.6),(1.8)):

$$\begin{aligned}
\frac{d\underline{b}}{dt}\cdot\underline{n} &= \frac{1}{||\alpha'(t)\times\alpha''(t)||}\,\alpha'(t)\times\alpha'''(t)\cdot\underline{n} = \\
&= \frac{1}{||\alpha'(t)\times\alpha''(t)||}\,\alpha'(t)\times\alpha'''(t)\cdot\underline{n} = \\
&= \frac{1}{||\alpha'(t)\times\alpha''(t)||}\,\alpha'(t)\times\alpha'''(t)\cdot\frac{1}{||\frac{d\underline{t}}{dt}||}\frac{d\underline{t}}{dt} = \\
&= \frac{1}{||\alpha'(t)\times\alpha''(t)||\,||\frac{d\underline{t}}{dt}||}\,\alpha'(t)\times\alpha'''(t)\cdot\frac{1}{||\alpha'(t)||}\alpha''(t) = \\
&= \frac{||\alpha'(t)||^2}{||\alpha'(t)\times\alpha''(t)||\,||^2}\,\alpha'(t)\times\alpha'''(t)\cdot\frac{1}{||\alpha'(t)||}\,\alpha''(t) = \\
&= \frac{||\alpha'(t)||}{||\alpha'(t)\times\alpha''(t)||\,||^2}\,\alpha'(t)\times\alpha'''(t)\cdot\alpha''(t).
\end{aligned}$$

we can conclude using the definition of ω_{32}. □

Example 1.11 *The straight line:* $\alpha(t) = (x_0+lt, y_0+mt, z_0+nt)$ *where* (x_0, y_0, z_0), $(l, m, n) \in \mathbb{R}^3$ *are constant. We have* $\alpha'(t) = (l, m, n)$ *and* $\alpha''(t) = (0, 0, 0)$ *hence the curve is not bi-regular and we cannot use the formula for the torsion of* α*. For the curvature we get* $\mathrm{k}(t) = 0$*.*

Proposition 1.8 *Let* $\alpha : I \to \mathbb{R}^3$ *be a regular curve. Then* $\mathrm{k}(t) = 0$ *if and only if* α *is a straight line.*

Proof: For one of the implications we can use Example 1.11. If α is regular we may assume that α is parametrized by arclenght. Then the norm of $\underline{t} = \alpha'(t)$ is constant and $\alpha''(t)$ is orthogonal to $\underline{t}$. Since the curvature vanishes we have $\alpha'(t) \times \alpha''(t) = 0$ and $\alpha''(t)$ is parallel to $\alpha'(t)$. Since $\alpha'(t) \neq 0$ this is only possible if $\alpha''(t) = 0$. Hence $\alpha'(t) = (l, m, n)$ for some constants l, m, n and if we integrate once more we get the equation of a line. the curvature □

Example 1.12 *The* $\alpha(t) = (x_0 + r\cos(t), y_0 + r\sin(t), 0)$ *be a circle in the* xy *plane in* $\mathbb{R}^3$*. We have* $\alpha'(t) = (-r\sin(t), r\cos(t), 0)$, $\alpha''(t) = (-r\cos(t), -r\sin(t), 0)$ *and* $\alpha'''(t) = -\alpha'(t)$*. Hence* $\alpha'(t) \times \alpha''(t) = (0, 0, r^2)$ *and* $||\alpha'(t)|| = r$ *and the curvature of* α *is constant:*

$$\mathrm{k}(t) = \frac{r^2}{r^3} = \frac{1}{r}.$$

To compute the torsion we can use the properties of the triple product:

$$\alpha'(t) \times \alpha'''(t) \cdot \alpha''(t) = -\alpha'(t) \times \alpha''(t) \cdot \alpha'''(t) = 0.$$

Hence

$$\tau(t) = 0.$$

Remark 1.2 *Let* $\alpha : I \to \mathbb{R}^3$ *be a bi-regular curve. Then the two planes spanned by* $\{\underline{t}, \underline{n}\}$ *and* $\{\alpha'(t), \alpha''(t)\}$ *coincide. In fact we have that* $\underline{t}$ *is parallel to* $\alpha'(t)$ *while* $\alpha''(t)$ *(see* (1.7)*) is a linear combination of* $\underline{t}$ *and* $\underline{n}$*. Since* $\underline{b}$ *is a unit vector perpendicular to that plane we must have*

$$\underline{b} = \pm \frac{1}{||\alpha'(t) \times \alpha''(t)||} \alpha'(t) \times \alpha''(t).$$

We have

$$\alpha'(t) \times \alpha''(t) = ||\alpha'(t)||\underline{t} \times (||\alpha'(t)||^2\ \mathrm{k}(t)\underline{n} + ||\alpha'(t)||'\ \underline{t}) = ||\alpha'(t)||^3\ \mathrm{k}(t)\ \underline{b}.$$

Since $\mathrm{k}(t) > 0$ *we have then*

$$\underline{b} = \pm\frac{1}{||\alpha'(t)\times\alpha''(t)||}\alpha'(t)\times\alpha''(t).$$

Proposition 1.9 *Let* $\alpha : I \to \mathbb{R}^3$ *be a bi-regular curve. Then* $\tau(t) = 0$ *if and only if* α *is a plane curve.*

Proof: Since the curve α is regular we may assume that α is parametrized by arclength. Suppose first that $\tau(t) = 0$. From the Frenet equation we get $\frac{d\underline{b}}{dt} = 0$ and $\underline{b}(t) = \underline{b}$ is constant. Let $t_0 \in I$ and

$$F(t) = \left((\alpha(t) - \alpha(t_0)) \cdot \underline{b}\right).$$

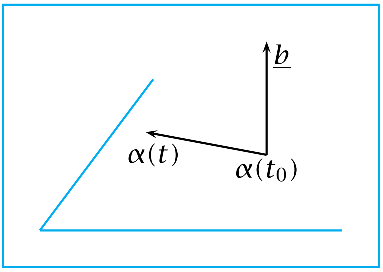

fig. 1.10

Then

$$\begin{aligned} F'(t) &= \frac{d}{dt}\left((\alpha(t) - \alpha(t_0)) \cdot \underline{b}\right) = \\ &= \alpha'(t) \cdot \underline{b} = \underline{t} \cdot \underline{b} = 0 \end{aligned}$$

so that $F(t) = F(t_0) = 0$. Hence, for every t, the vector $\alpha(t) - \alpha(t_0)$ lies in a plane orthogonal to $\underline{b}$ (Fig. 1.10). In particular α is a plane curve. Suppose now that, for every t, $\alpha(t)$ lies in some plane. Denote by $\underline{v}$ a vector orthogonal to that plane. Then

$$0 = \frac{d}{dt}\left((\alpha(t) - \alpha(t_0)) \cdot \underline{v}\right) = \alpha'(t) \cdot \underline{v}$$

taking another derivative we have $\alpha''(t) \cdot \underline{v} = 0$. Hence $\underline{v}$ is orthogonal to the plane spanned by $\{\alpha'(t), \alpha''(t)\}$ but (see Remark 1.2) this implies that $\underline{v}$ is parallel to $\underline{b}(t)$. Hence $\underline{b}(t) = \underline{b}$ is constant. From the Frenet equations it follows that $\tau(t) = 0$. □

Example 1.13 *If the curve is not bi-regular a singularity in the Frenet frame may occur at the points where* $\mathrm{k} = 0$*. Let*

$$\alpha(t) = \begin{cases} (t, 0, e^{-\frac{1}{t^2}}) & t < 0 \\ (0,0,0) & t = 0 \\ (t, e^{-\frac{1}{t^2}},) & t > 0 \end{cases}$$

The curve is not bi-regular at $t = 0$*, the torsion is zero in an open dense set but the curve lies in the* xz *plane for* $t < 0$ *and in the* xy *plane for* $t > 0$.

The fact that the curvature is constant characterizes the circle among the plane curves:

Proposition 1.10 *Let $\alpha : I \to \mathbb{R}^3$ be a bi-regular plane curve such that $\mathrm{k}(t) = k$ is constant. Then the trace of α is a portion of a circle.*

Proof: Assume that the curve is parametrized by arclength and let $\underline{n}$ be the normal vector in the Frenet frame. Let

$$F(t) = \alpha(t) + \frac{1}{k}\underline{n}.$$

Since α is a plane curve the torsion τ is zero and we have

$$F'(T) = \underline{t} + \frac{1}{k}(-k\underline{t}) = 0.$$

Hence $F(t) = p$ for some point $p \in \mathbb{R}^3$, then

$$||\alpha(t) - p|| = ||\frac{1}{k}\underline{n}|| = \frac{1}{k}$$

and $\alpha(t)$ is a point of the circle of radius $\frac{1}{k}$ centered in p. □

Example 1.14 *The $\alpha(t) = (r\cos(t), r\sin(t), at)$ be a cylindrical helix in $\mathbb{R}^3$. This is a bi-regular curve and we have $\alpha'(t) = (-r\sin(t), r\cos(t), a)$, $\alpha''(t) = (-r\cos(t), -r\sin(t), 0)$ and $\alpha'''(t) = (r\sin(t), -r\cos(t), 0)$. Hence $\alpha'(t)\times\alpha''(t) = (ar\sin(t), -ar\cos(t), r^2)$ and $||\alpha'(t)|| = \sqrt{a^2 + r^2}$. The curvature of α is constant:*

$$\mathrm{k}(t) = \frac{r}{r^2 + a^2} = \frac{1}{r}.$$

Since $\alpha'(t) \times \alpha''(t) \cdot \alpha'''(t) = -ar^2$. the torsion of α is also constant

$$\tau(t) = -\frac{a}{r^2 + a^2}.$$

For the helix we have that $\underline{v} = (1, 0, 0)$ is a constant vector such that $\underline{v} \cdot \alpha'(t) = a$ is also constant. We can use this fact to generalize the definition:

Definition 1.9 *A bi-regular curve $\alpha : I \to \mathbb{R}^3$ is a generalized helix if the tangent vector $\alpha'(t)$ makes a constant angle with a fixed unit vector $\underline{v}$, i.e. $\underline{t} \cdot \underline{v} = \cos(\theta)$ is constant.*

Proposition 1.11 *A bi-regular curve $\alpha : I \to \mathbb{R}^3$ is a generalized helix if and only if the ratio τ / k is constant.*

Proof: Suppose first that $\alpha(t) : I \to \mathbb{R}^3$ is a generalized helix parametrized by arclength. Then $\underline{t} \cdot \underline{v} = \cos(\theta)$ and, taking the derivative, $\mathrm{k}(t)\underline{n} \cdot \underline{v} = 0$. Since the curvature of α is positive we have

$$\underline{n} \cdot \underline{v} = 0 \tag{1.9}$$

and $\underline{v}$ belongs to the plane spanned by $\underline{t}$ and $\underline{b}$. Since the angle between $\underline{v}$ and $\underline{t}$ is θ we have that $\underline{b} \cdot \underline{v} = \pm\sin(\theta)$. Taking the derivative of (1.9) we obtain

$$0 = \frac{d}{dt}(\underline{n} \cdot \underline{v}) = -(\mathrm{k}(t)\,\underline{t} + \tau(t)\,\underline{b}) \cdot \underline{v} = -\cos(\theta)\,\mathrm{k}(t) \pm \sin(\theta)\tau(t).$$

It follows that $\tau/\mathrm{k} = \pm\cot(\theta)$ is constant.

For the converse, let θ be an angle such that $\cot(\theta) = \tau/\mathrm{k}$. Let

$$\underline{v} = \cos(\theta)\,\underline{t} - \sin(\theta)\,\underline{b}.$$

Then

$$\frac{d}{dt}\underline{v} = \cos(\theta)\,\mathrm{k}(t)\,\underline{n} - \sin(\theta)\tau(t)\,\underline{n} = \mathrm{k}(t)(\cos(\theta) - \sin(\theta)\frac{\tau(t)}{\mathrm{k}(t)})\,\underline{n} = 0.$$

So that $\underline{v}$ is a constant unit vector and $\underline{t} \cdot \underline{v} = \cos(\theta)$ is constant. □

We conclude this section with the so called local canonical form. Assume that $\alpha : [0, L] \to \mathbb{R}^3$ is a C^3 bi-regular curve parametrized by arclength such that $\alpha(t_0) = (0,0,0)$. We can write

$$\alpha(t) = \alpha(t_0) + t\alpha'(t_0) + \frac{t^2}{2}\alpha''(t_0) + \frac{t^3}{6}\alpha'''(t_0) + o((t-t_0)^3)$$

where, for t close to t_0,

$$\lim_{t\to t_0} \frac{o((t-t_0)^3)}{(t-t_0)^3} = 0.$$

We have $\alpha(t_0) = 0$ and

$$\alpha'(t_0) = \underline{t}(t_0),$$

$$\alpha''(t_0) = \mathrm{k}(t_0)\,\underline{n}(t_0),$$

$$\alpha'''(t_0) = \mathrm{k}'(t_0)\,\underline{n}(t_0) - \mathrm{k}(t_0)\left(\mathrm{k}(t_0)\,\underline{t} + \tau(t_0)\,\underline{b}\right).$$

Hence

$$\begin{aligned} \alpha(t) \quad &= \quad \left(t - \frac{1}{6}\mathrm{k}(t_0)^2\, t^3\right)\underline{t} + \left(\frac{1}{2}\mathrm{k}(t_0)\,t^2 + \frac{1}{6}\mathrm{k}'(t_0)\,t^3\right)\underline{n} \\ &- \quad \left(\frac{1}{6}\mathrm{k}(t_0)\tau(t_0)\,t^3\right)\underline{b} + o(t^3). \end{aligned}$$

Using this formula it is possible to derive some information about the local behavior of the curve near $t = t_0$, in particular with respect to some planes:

Definition 1.10 *Let $\alpha : I \to \mathbb{R}^3$ be a bi-regular curve. The plane spanned by*

(i) *$\underline{n}$ and $\underline{b}$ is the normal plane at $\alpha(t)$*
(ii) *$\underline{t}$ and $\underline{b}$ is the rectifying plane at $\alpha(t)$*
(iii) *$\underline{t}$ and $\underline{n}$ is the osculating plane at $\alpha(t)$.*

From the local canonical form we see that the curve α, for t close enough to t_0

(i) stays on one side of the rectifying plane (since k > 0)
(ii) crosses the osculating plane in the direction of $\underline{b}$ (resp $-\underline{b}$) if the torsion is negative (resp. positive)

Global results

We will now prove three theorems that show the importance of the curvature and the torsion in the theory of space curves. First note that if $x(t)$ is a C^1 family of vectors in $\mathbb{R}^3$ depending on time and $f_A : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear map then, since A does not depend on t,

$$\frac{d}{dt} f_A(x(t)) = \frac{d}{dt} A \cdot x(t) = A \cdot x'(t) = f_A(x'(t)).$$

Theorem 1.1 *Let $\alpha : I \to \mathbb{R}^3$ be a bi-regular curve and let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be an orientation preserving isometry. Let $\beta(t) = f(\alpha(t))$. Then β and α have the same curvature and torsion at the corresponding points.*

Proof: Any isometry is the composition of a linear map and a translation. If f is a translation then the result is obvious since the components of α and β only differ by a constant and the curvature and the torsion are computed by taking derivatives of α and β. Hence it is enough to consider the case $f = f_A$ where $A \in O(3)$ is an orthogonal matrix with positive determinant. We assume that α is parametrized by arclength. Then

$$f_A(\underline{t}_\alpha) = f_A(\alpha'(t)) = \frac{d}{dt} f_A(\alpha(t)) = \frac{d}{dt}\beta(t) = \beta'(t).$$

Since f_A is an isometry $\beta'(t)$ has unit length and

$$f_A(\underline{t}_\alpha) = \underline{t}_\beta.$$

Similarly

$$f_A(\frac{d}{dt}\underline{t}_\alpha) = \frac{d}{dt}\underline{t}_\beta$$

since the two vectors have the same norm we have

$$f_A(\underline{n}_\alpha) = \underline{n}_\beta.$$

The map f_A is an orientation preserving isometry. Hence

$$\{f_A(\underline{t}_\alpha), f_A(\underline{n}_\alpha), f_A(\underline{b}_\beta)\} = \{\underline{t}_\beta, \underline{n}_\beta, f_A(\underline{b}_\alpha)\}$$

is a positive orthonormal basis. It follows that $f_A(\underline{b}_\alpha) = \underline{b}_\beta$ and f_A maps the Frenet frame of α into the Frenet frame of β at every point. We have

$$\frac{d\underline{t}_\beta}{dt} = \frac{d}{dt} f_A(\underline{t}_\alpha) = f_A\left(\frac{d\underline{t}_\alpha}{dt})\right) = f_A(\mathrm{k}_\alpha\, \underline{n}_\alpha) = \mathrm{k}_\alpha\, \underline{n}_\beta.$$

Comparing with the Frenet's equations for β we obtain $\mathrm{k}_\alpha = \mathrm{k}_\beta$.

$$\frac{d\underline{b}_\beta}{dt} = \frac{d}{dt} f_A(\underline{b}_\alpha) = f_A\left(\frac{d\underline{b}_\alpha}{dt}\right) = f_A(\tau_\alpha\, \underline{n}_\alpha) = \tau_\alpha\, \underline{n}_\beta.$$

Comparing with the Frenet's equations for β we obtain $\tau_\alpha = \tau_\beta$. □

Hence curvature and torsion are invariant under isometries. Note that if f_A is orientation reversing we still have $f_A(\underline{t}_\alpha) = \underline{t}_\beta$ and $f_A(\underline{n}_\alpha) = \underline{n}_\beta$ but $f_A(\underline{b}_\alpha) = -\underline{b}_\beta$ hence the curvature of β is still the same as the curvature of α but the torsion of β and the torsion of α differ by the sign.

We prove a 'converse' to the previous theorem:

Theorem 1.2 *Let $\alpha, \beta : I \to \mathbb{R}^3$ be two bi-regular curves such that $||\alpha'(t)|| = ||\beta'(t)||$, $\mathrm{k}_\alpha(t) = \mathrm{k}_\beta(t)$ and $\tau_\alpha(t) = \tau_\beta(t)$ for every $t \in I$. Then there exists an orientation preserving isometry $f : \mathbb{R}^3 \to \mathbb{R}^3$ such that $\beta(t) = f(\alpha(t))$.*

Proof: We fix $t_0 \in I$. Then $\{\underline{t}_\alpha(t_0), \underline{n}_\alpha(t_0), \underline{b}_\alpha(t_0)\}$ $\{\underline{t}_\beta(t_0), \underline{n}_\beta(t_0), \underline{b}_\beta(t_0)\}$ are two positive orthonormal bases of $\mathbb{R}^3$ and we can find an orientation preserving isometry $f_A : \mathbb{R}^3 \to \mathbb{R}^3$ that maps one onto the other. We, for every $t \in I$, three families of vectors depending on t:

$$x(t) = f_A(\underline{t}_\alpha(t)), \quad y(t) = f_A(\underline{n}_\alpha(t)), \quad z(t) = f_A(\underline{b}_\alpha(t).$$

We have

$$\begin{aligned} \frac{dx(t)}{dt} &= \frac{d}{dt} f_A(\underline{t}_\alpha(t)) = f_A(\frac{d}{dt}\underline{t}_\alpha(t)) = f_A(\mathrm{k}_\alpha(t)\, ||\alpha'(t)||\, \underline{n}_\alpha(t)) = \\ &= \mathrm{k}_\alpha(t)\, ||\alpha'(t)||\, f(\underline{n}_\alpha(t)) = \mathrm{k}_\beta(t)\, ||\beta'(t)||\, y(t) \end{aligned}$$

$$\begin{aligned} \frac{dy(t)}{dt} &= \frac{d}{dt} f_A(\underline{n}_\alpha(t)) = f_A(\frac{d}{dt}\underline{n}_\alpha(t)) = \\ &= f_A(-\mathrm{k}_\alpha(t)\, ||\alpha'(t)||\, \underline{t}_\alpha(t) - \tau_\alpha(t)\, ||\alpha'(t)||\, \underline{b}_\alpha(t)) = \end{aligned}$$

$$\begin{aligned} &= -\mathrm{k}_\alpha(t)\,||\alpha'(t)||\,f_A(\underline{t}_\alpha(t)) - \tau_\alpha(t)\,||\alpha'(t)||\,f_A(\underline{b}_\alpha(t)) = \\ &= -\mathrm{k}_\beta(t)\,||\beta'(t)||\,x(t) - \tau_\beta(t)\,||\beta'(t)||\,z(t) \end{aligned}$$

$$\begin{aligned} \frac{dz(t)}{dt} &= \frac{d}{dt} f_A(\underline{b}_\alpha(t)) = f_A(\frac{d}{dt}\underline{b}_\alpha(t)) = f_A(\tau_\alpha(t)\,||\alpha'(t)||\,\underline{n}_\alpha(t)) = \\ &= \tau_\alpha(t)\,||\alpha'(t)||\,f(\underline{n}_\alpha(t)) = \tau_\beta(t)\,||\beta'(t)||\,y(t) \end{aligned}$$

Hence $\{x(t), y(t), z(t)\}$ satisfies a system of linear vector ODE's with initial conditions $x(t_0) = \underline{t}_\beta(t_0)$,$y(t_0) = \underline{n}_\beta(t_0)$, $z(t_0) = \underline{b}_\beta(t_0)$. By the Cauchy theorem $x(t), y(t), z(t)$ is the unique solution of the system and it is defined in the whole interval I. On the other hand we have that $\underline{t}_\beta(t), \underline{n}_\beta(t), \underline{n}_\beta(t)$ is another solution of the same equations with the same initial condition. We can conclude that

$$\underline{t}_\beta(t) = x(t) = f_A(\underline{t}_\alpha(t)), \quad \underline{n}_\beta(t) = y(t) = f_A(\underline{n}_\alpha(t)), \quad \underline{n}_\beta(t) = z(t) = f_A(\underline{b}_\alpha(t)$$

and the isometry f_A maps the Frenet frame of α into the Frenet frame of β for any t. In particular the tangent vector $\alpha'(t)$ to α is mapped into the tangent vector $\beta'(t)$ to β. Let

$$\gamma(t) = f_A(\alpha(t)) + \beta(t_0) - f_A(\alpha(t_0)).$$

Then γ is obtained by applying to α the isometry f_A and a translation defined by the vector $\beta(t_0) - f_A(\alpha(t_0))$. Hence γ and α differ by an orientation preserving isometry. We have

$$\gamma'(t) = \frac{d}{dt} f_A(\alpha(t)) = f_A(\alpha'(t)) = \beta'(t)$$

with $\gamma(t_0) = \beta(t_0)$. Hence $\gamma(t) = \beta(t)$. □

Corollary 1.3 *A bi-regular curve $\alpha : I \to \mathbb{R}^3$ is an helix if and only if the curvature and the torsion are constant.*

Proof: Given α it is possible to find an helix that has the same curvature and torsion. Hence the two curves are isometric and α is an helix. □

This shows that once the curvature and the torsion of a bi-regular curve are given then the curve is uniquely determined up to isometries. The next result shows that it is possible to prescribe the curvature and the torsion arbitrarly:

Theorem 1.3 *Let* $\mathrm{k}, \tau : I \to \mathbb{R}$ *be two smooth functions such that* $\mathrm{k}(t) > 0$ *for every* $t \in I$. *Given* $t_0 \in I$, $p \in \mathbb{R}^3$ *and a positive orthonormal basis* $\{v_1, v_2, v_3\}$ *of* $\mathbb{R}^3$ *there exists a bi-regular curve* $\alpha : I \to \mathbb{R}^3$ *such that*

(i) $\alpha(t_0) = p$

(ii) $||\alpha'(t)|| = 1$ *for every* $t \in I$

(iii) $\underline{t}(t_0) = v_1$, $\underline{n}(t_0) = v_2$, $\underline{b}(t_0) = v_3$

(iv) $\mathrm{k}_\alpha(t) = \mathrm{k}(t)$, $\tau_\alpha(t) = \tau(t)$ *for every* $t \in I$

Proof: The first three parts of the statement are in some sense obvious, we can always parametrize α with the arclength, apply an isometry f_A such that the Frenet frame of α at $t = t_0$ coincides with $\{v_1, v_2, v_3\}$ and then a translation that moves $\alpha(t_0)$ in p. The main thing here is the statement about the curvature and the torsion.

Let us consider the system of vector ODE's in the unknowns $x(t), y(t), z(t)$

$$\begin{cases} \frac{dx(t)}{dt} = \mathrm{k}(t)\, y(t) \\ \frac{dy(t)}{dt} = -\mathrm{k}(t)\, x(t) - \tau(t)\, z(t) \\ \frac{dz(t)}{dt} = \tau(t)\, y(t) \end{cases}$$

with initial conditions $x(t_0) = v_1$,$y(t_0) = v_2$, $z(t_0) = v_3$ This is a system of linear ODE's and by the Cauchy theorem there is a unique solution $x(t), y(t), z(t)$ defined in the whole interval I. Let

$$\alpha(t) = p + \int_{t_0}^{t} x(t)\, dt.$$

It is clear that $\alpha(t_0) = p$. Now we want to prove that α is parametrized by arclength. Let

$$\Omega(t) = \begin{pmatrix} 0 & \mathrm{k}(t) & 0 \\ -\mathrm{k}(t) & 0 & -\tau(t) \\ 0 & \tau(t) & 0 \end{pmatrix}, \qquad A(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}.$$

Note that $A(t)$ is a 3×3 matrix whose rows are determined by the components of the vectors $x(t)$, $y(t)$, $z(t)$. Then we can write the system in matrix form as

$$A'(t) = \Omega(t) \cdot A(t).$$

We know that $A(t_0)$ is an orthogonal matrix with positive determinant (in fact for $t = t_0$ the rows of A coincide with v_1, v_2, v_3). If we prove that $A(t) \in O(3)$ for all t then $x(t)$, $y(t)$, $z(t)$ will be an orthonormal basis of $\mathbb{R}^3$, in particular

$||\alpha'(t)|| = ||x(t)|| = 1$. Let $B = A^t \cdot A$. Then, using the fact that Ω is skew-symmetric

$$\begin{aligned} B' &= \left(A^t \cdot A\right)' = A^{t\prime} \cdot A + A^t \cdot A' = A'^t \cdot A + A^t \cdot A' = \\ &= (\Omega \cdot A)^t \cdot A + A^t \cdot (\Omega \cdot A) = A^t \cdot \Omega^t \cdot A + A^t \cdot \Omega \cdot A = \\ &= -A^t \cdot \Omega \cdot A + A^t \cdot \Omega \cdot A = 0 \end{aligned}$$

Hence $B(t) = B(t_0) = I$ and $A(t)$ is orthogonal. Note that the determinant $\det(A(t))$ is a continuous function of the entries of $A(t)$. Since this determinant can only be equal to ± 1 and it is equal to 1 at $t = t_0$ we can conclude that $A(t)$ is always orientation preserving. Now we know that $\alpha(t)$ is parametrized by arclength and $\alpha'(t) = x(t) = \underline{t}$. Hence

$$\frac{d\underline{t}}{dt} = \frac{dx(t)}{dt} = \mathrm{k}(t)y(t)$$

since $\mathrm{k}(t) > 0$ and $y(t)$ is a unit vector it follows that $y(t) = \underline{n}$ and $\mathrm{k}_\alpha = \mathrm{k}$. Since we have positively oriented frames we may conclude that $\underline{b} = z(t)$ and

$$\frac{d\underline{b}}{dt} = \frac{dz(t)}{dt} = \tau(t)y(t) = \tau(t)\underline{n}.$$

In particular $\tau_\alpha = \tau$. $\square$

Chapter 2

Surfaces

Surfaces

A regular surface in $\mathbb{R}^3$ is a subset of $\mathbb{R}^3$. Remember that a curve is a map, not a set. Hence we are taking a different approach here. The reason is that if we consider sets instead of maps it is easier to prove global results. Our first goal is to use the tools of calculus to study the surfaces, the first tool is the differential of a function:

Definition 2.1 *Let $U \subset \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}^m$ be a smooth map. If $p \in U$ then the differential of f at p is the linear map $df(p) : \mathbb{R}^n \to \mathbb{R}^m$ associated, with respect to the standard basis, to the matrix*

$$df(p) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

Now we give the definition of regular surface. The idea is the we smoothly deform and patch together open subsets of $\mathbb{R}^2$. There is a technical part in the definition that is necessary to avoid singularities like corners or cuspidal points or the lack of a tangent plane.

Definition 2.2 *A regular surface is a subset $S \subset \mathbb{R}^3$ such that for every point $p \in S$ there exists a neighborhood V of p in $\mathbb{R}^3$ and a surjective map $X : U \to V \cap S$ defined in an open set $U \subset \mathbb{R}^2$ such that*

(i) *X is differentiable, i.e. $X(u,v) = (x(u,v), y(u,v), z(u,v))$ where x, y, z are smooth functions $U \to \mathbb{R}$.*

(ii) *X is an homeomorphism, i.e. X has a continuous inverse $X^{-1} : V \cap S \to U$.*

(iii) *The differential $dX(q) : \mathbb{R}^2 \to \mathbb{R}^3$ is injective at every point $q \in U$.*

The map X is called local parametrization near p

Recall that a function $V \cap S \to \mathbb{R}^2$ of $\mathbb{R}^3$ is continuous if it is the restriction of a continuous function defined in an open set containing $V \cap S$. If $X : \mathbb{R}^2 \to \mathbb{R}^3$,

$X(u,v) = (x(u,v), y(u,v), z(u,v))$ is differentiable then the differential of X at q is the linear map $\mathbb{R}^2 \to \mathbb{R}^3$ defined, with respect to the standard bases of $\mathbb{R}^2$ and $\mathbb{R}^3$, by the matrix

$$dX(q) = \begin{pmatrix} \frac{\partial x}{\partial u}(q) & \frac{\partial x}{\partial v}(q) \\ \frac{\partial y}{\partial u}(q) & \frac{\partial y}{\partial v}(q) \\ \frac{\partial z}{\partial u}(q) & \frac{\partial z}{\partial v}(q) \end{pmatrix}$$

the differential is injective if and only if $dX(q)$ has rank 2, i.e. the two column vectors in $dX(q)$ are linearly independent. This is equivalent to the fact that at least one of the determinants of the 2×2 minors of $dX(q)$

$$\begin{pmatrix} \frac{\partial x}{\partial u}(q) & \frac{\partial x}{\partial v}(q) \\ \frac{\partial y}{\partial u}(q) & \frac{\partial y}{\partial v}(q) \end{pmatrix}, \qquad = \begin{pmatrix} \frac{\partial y}{\partial u}(q) & \frac{\partial y}{\partial v}(q) \\ \frac{\partial z}{\partial u}(q) & \frac{\partial z}{\partial v}(q) \end{pmatrix}, \qquad \begin{pmatrix} \frac{\partial x}{\partial u}(q) & \frac{\partial x}{\partial v}(q) \\ \frac{\partial z}{\partial u}(q) & \frac{\partial z}{\partial v}(q) \end{pmatrix}$$

is different from zero.

Example 2.1 *Let $f : U \to \mathbb{R}$ be a smooth function, where $U \subset \mathbb{R}^2$ is an open set. Let*

$$S = \{(x,y,z) \in \mathbb{R}^3 : z = f(x,y)\}$$

be the graph of f. Consider the function $X : U \to \mathbb{R}^3$

$$X(u,v) = (u, v, f(u,v)).$$

By definition, $X(U) = S$. The function X is differentiable, injective and surjective on S. If $(x,y,z) \in S$ then $X^{-1}(x,y,z) = (x,y)$, a continuous map. Moreover

$$dX(u,v) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix}$$

and this matrix has constant rank 2. Hence S is a regular surface, covered by just one coordinate neighborhood. This shows that planes (for $f(x,y) = ax + by + c$), the paraboloid (for $f(x,y) = x^2 + y^2$), the hyperboloid (for $f(x,y) = x^2 - y^2$) are regular surfaces.

Example 2.2 *Fix $r > 0$ and let $S^2(r) = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = r^2\}$ be the sphere of radius r centered in the origin. Let $U = \{(u,v) \in \mathbb{R}^2 : u^2 + v^2 < r^2\}$ and consider the map*

$$X_1(u,v) = (u, v, \sqrt{r^2 - u^2 - v^2}).$$

The image of X_1 is an open subset of $S^2(r)$ (the upper hemisphere) and, since we are showing that it is the graph of the smooth function $f(x,y) = \sqrt{r^2 - u^2 - v^2}$

we have that the definition of regular surface is satisfied by the upper hemisphere. If we consider

$$X_2(u,v) = (u, v, -\sqrt{r^2 - u^2 - v^2})$$

then we have that $S^2(r)$ minus the circle lying in the xy plane, satisfies the definition of regular surface. To prove that $S^2(r)$ is a regular surface we have to cover it with four more parametrization:

$$X_3(u,v) = (u, \sqrt{r^2 - u^2 - v^2}, v), \qquad X_4(u,v) = (u, -\sqrt{r^2 - u^2 - v^2}, v)$$

that cover the points with $y > 0$ and $y < 0$, and (we still have two points $(\pm 1, 0, 0)$ missing)

$$X_5(u,v) = (\sqrt{r^2 - u^2 - v^2}, u, v), \qquad X_4(u,v) = (-\sqrt{r^2 - u^2 - v^2}, u, v).$$

Hence, locally, the sphere is always the graph of a function defined in one of the coordinate planes. This is a general fact, to show it we will use a fundamental tool of the multivariate calculus:

Theorem 2.1 *[Inverse function theorem] Let $U \subset \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}^n$ be a smooth map. If $dF(p)$ is an isomorphism then f is a local diffeomorphism.*

In particular it is possible to find an open neighborhood U' of p such that $f : U' \to f(U')$ is a diffeomorphism (i.e. it is smooth with a smooth inverse). In the case of a function of one variable this just means that a differentiable function $f : \mathbb{R} \to \mathbb{R}$ such that $f'(x_0) \neq 0$ (i.e. a function that is monotone close to x_0) is invertible and has a smooth inverse.

Example 2.3 *The function $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by*

$$f(x,y) = (x^2 - y^2, 2xy)$$

is not injective (we have $(f(-x,-y) = f(x,y))$ but

$$df(x,y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$$

has determinant $4(x^2 + y^2)$ hence every point different from the origin has a neighborhood U such that f restricted to U has a smooth inverse.

Proposition 2.1 *Let S be a regular surface. If $p \in S$ then there is a neighborhood V of p in S such that V is the graph of a function.*

Proof: Let $p \in S$ and let $X : U \to S$ be a local parametrization and let $q = X^{-1}(p)$. Assume, for example, that

$$\det \begin{pmatrix} \frac{\partial x}{\partial u}(q) & \frac{\partial x}{\partial v}(q) \\ \frac{\partial y}{\partial u}(q) & \frac{\partial y}{\partial v}(q) \end{pmatrix} \neq 0$$

(otherwise we use another pair of coordinates). And define the function $\phi : U \to \mathbb{R}^2$, $\phi(u,v) = \pi \circ X = (x(u,v), y(u,v))$, where π is the projection on the xy coordinate plane in $\mathbb{R}^3$. Then $d\phi(q)$ is an isomorphism and we have a neighborhood U' of q such that ϕ, restricted to U', has a smooth inverse: $\psi(x,y) = (u(x,y), v(x,y))$. Then $\tilde{X} = X \circ \psi : \phi(U') \to S$ is a local parametrization in a neighborhood of p and it is given by

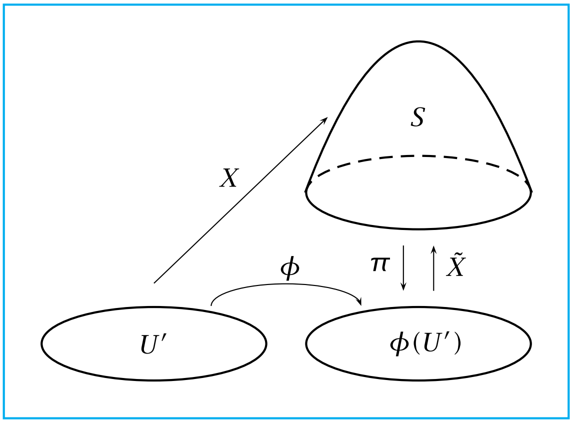

fig. 2.1

$$\begin{aligned} \tilde{X}(x,y) &= (x(u(x,y), v(x,y), y(u(x,y), v(x,y)), z(u(x,y), v(x,y)) = \\ &= (x, y, z(u(x,y), v(x,y)) \end{aligned}$$

i.e. S is locally the graph of a function. □

In general it is not easy to prove directly that a function X is an homeomorphism. We would like to have tools that help us constructing examples without having to check it every time. We need another general result:

Definition 2.3 *Let $f : V \to \mathbb{R}$ where V is an open subset of $\mathbb{R}^n$. A number $a \in \mathbb{R}$ is a regular value for f if $f^{-1}(a) \neq \emptyset$ and, for every $p \in f^{-1}(a)$, the gradient $\nabla f(p)$ is non zero.*

Theorem 2.2 *[Implicit function theorem] Let $V \subset \mathbb{R}^n$ be an open subset and $f : V \to \mathbb{R}$ be a smooth map. If $a \in \mathbb{R}$ is a regular value for f. If $p = (x_1, \dots, x_n) \in f^{-1}(a)$ and $\frac{\partial f}{\partial x_i}(p) \neq 0$ then there exists a neighborhood W of p, a neighborhood U of $\pi_i(p) = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in \mathbb{R}^{n-1}$ and a smooth function $g : U \to \mathbb{R}^n$ such that*

$$f^{-1}(a) \cap W = \{(x_1, \dots, x_{i-1}, g(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n), x_{i+1}, \dots, x_n)\}$$

for $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in U$.

Proposition 2.2 *Let $f : V \subset \mathbb{R}^3 \to \mathbb{R}$ be a smooth function. If $a \in \mathbb{R}$ is a regular value for f then $f^{-1}(a)$ is a regular surface in $\mathbb{R}^3$.*

Proof: This is an immediate consequence of the implicit function theorem. In fact we have that, locally, $f^{-1}(a)$ is the graph of a smooth function, i.e. a regular surface. □

Example 2.4 *The sphere $S^2(r)$ can be defined as $f^{-1}(r^2)$, where $f(x,y,z) = x^2+y^2+z^2$. The gradient of f is the vector*

$$\nabla f = 2(x,y,z)$$

and if $p\S^2(r)$, with $r>0$, this vector is nonzero. In fact we have seen before that the points $p=(x,y,z)$ with $x>0$ belong to the graph of the function X_1...

Example 2.5 *The torus T^2 is the surface*

$$T^2 = \{(x,y,z) \in \mathbb{R}^3 : (\sqrt{x^2+y^2}-a)^2 + z^2 = r^2\}.$$

generated by a circle of radius r in the xz plane, centered at $(a,0,0)$ (where $a>r>0$), rotated around the z axis. To show that T^2 is a regular surface we consider the function

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

and we have to prove that r^2 is a regular value for f. In fact f is a smooth function and at least one of the partial derivatives

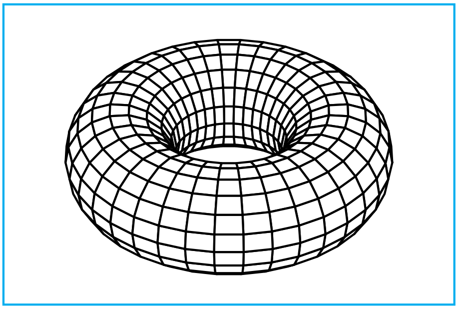

fig. 2.2

$$\frac{\partial f}{\partial x} = \frac{2x(\sqrt{x^2+y^2}-a)}{\sqrt{x^2+y^2}}, \qquad \frac{\partial f}{\partial y} = \frac{2y(\sqrt{x^2+y^2}-a)}{\sqrt{x^2+y^2}} \qquad \frac{\partial f}{\partial z} = 2z$$

does not vanish at $f^{-1}(r^2)$. Hence r^2 is a regular value for f and T^2 is a regular surface.

Example 2.6 *The torus is a special case of rotational surface. Let $\alpha : [a,b] \to \mathbb{R}^3$, $\alpha(t) = (f(t),0,g(t))$ be a regular curve with no self-intersections (the curve is allowed to be closed i.e. $\alpha(a)=\alpha(b)$) such that $f(t)>0$ (i.e. the curve does not intersect the z axis). Then we define*

$$S = \{(f(v)\cos(u), f(v)\sin(u), g(v) \in \mathbb{R}^3 : v \in (a,b), u \in \mathbb{R}\}.$$

It is possible to cover S with two coordinate neighborhoods (with $u \in (0, 2\pi)$ and $(\pi, 3\pi)$ respectively) and prove that S is a regular surface. We will not give the details. Note that, from the fact that α is assumed to be injective, it follows that (u, v) can be uniquely determined once $(x, y, z) \in S$ is given.

If $\alpha(t) = (a\cosh(t), 0, at)$ where a is a positive constant, the resulting surface is called catenoid. The curve α is called catenary and it is related to the shape taken by a chain hanging with two fixed ends.

Example 2.7 *The helicoid is the surface obtained by considering all the lines joining a point of a circular helix*

$$(a\cos(t), a\sin(t), bt)$$

with the point $(0, 0, bt)$ of the z axis having the same height:

$$S = \{(av\cos(u), av\sin(u), bu\}$$

Where $u \in (0, 2\pi)$ and $v \in \mathbb{R}$. it is possibile to prove that S is a regular surface.

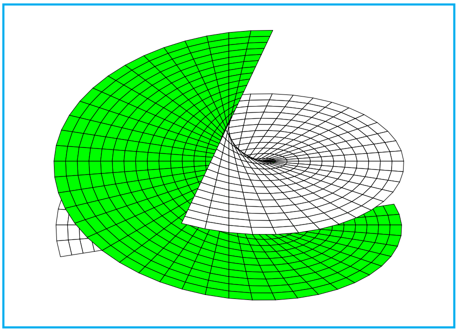

fig. 2.3

The tangent space

The main tool we used studying the theory of curves was the Frenet frame. At each point of a bi-regular curve α in $\mathbb{R}^3$ we have a positive orthonormal basis of $\mathbb{R}^3$. In other words we attach at every point of α a linear space spanned by these vectors and we study how the basis of this space changes along the curve. We want to do something similar for the surfaces. The first step will be the definition of tangent space. In the case of a curve α this was the space spanned by α' and it was clear that this was independent on the chosen parametrization of α. Here we have to work a little more..

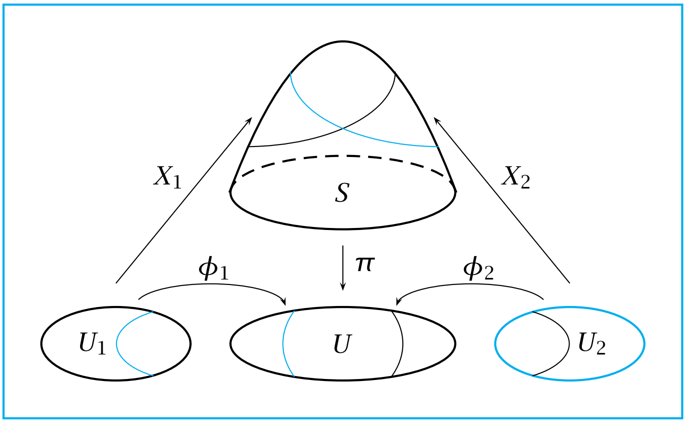

fig. 2.4

Suppose that, in a neighborhood of a point $p \in S$ we have two different parametrizations $X_1 : U_1 \to S$ and $X_2 : U_2 \to S$. Then $W = X_1(U_1) \cap X_2(U_2) \neq \emptyset$ and we have a map

$$X_{12} = X_2^{-1} \circ X_1 : X_1^{-1}(W) \to X_2(W).$$

From the definition of S it follows that X_{12} is an homeomorphism. We know that, locally, S is the graph of a smooth function f defined, for example, in an open subset of the xy plane. We have already used the implicit function theorem to show that the maps $\phi_1 = \pi \circ X_1$ and $\phi_2 = \pi \circ X_2$ are local diffeomorphisms close to $X_1^{-1}(p)$ and $X_2^{-1}(p)$ respectively. It follows that $X_{12} = \phi_2^{-1} \circ \phi_1$ is a diffeomorphism close to $X_1^{-1}(p)$.

Proposition 2.3 *Let S be a regular surface, $p \in S$ and $X_1 : U_1 \to S$, $X_2 : U_2 \to S$ two local parametrizations near p. Then the map*

$$X_2^{-1} \circ X_1 : X_1^{-1}(X_1(U_1) \cap X_2(U_2)) \to X_2^{-1}(X_1(U_1) \cap X_2(U_2))$$

is a diffeomorphism.

This is an important fact as it shows that 'smooth' objects defined in S using some parametrization remain 'smooth' with respect to any other local parametrization:

Definition 2.4 *Let S be a regular surface. Then $f : S \to \mathbb{R}^n$ is differentiable if $f \circ X : U \to \mathbb{R}^n$ is a smooth map for every local parametrization $X : U \to S$ of S.*

It follows that, in order to check that a function $f : S \to \mathbb{R}^n$ is smooth in a neighborhood of a point $p \in S$, it is enough to take one local parametrization X_1

near p. If X_2 is a different parametrization then $f \circ X_2 = f \circ X_1 \circ X_{12}^{-1}$ is smooth if and only if $f \circ X_2$ is smooth.

Remark 2.1 *Sometimes when $f : S \to \mathbb{R}^n$ is a smooth map, instead of writing $f(p)$ we will use the notation $f(u, v)$ justified by the fact that, given $f \circ X^{-1}$, f is uniquely determined.*

Definition 2.5 *Let S_1, S_2 be regular surfaces. Then $f : S_1 \to S_2$ is differentiable map if $X_2^{-1} \circ f \circ X_1 : U_1 \to U_2$ is a smooth map for every local parametrizations $X_1 : U_1 \to S_1$ and $X_2 : U_2 \to S_2$ such that $f(X_1(U_1)) \subset X_2(U_2)$. If f has a differentiable inverse then we say that f is a diffeomorphism.*

From the point of view of differential geometry two diffeomorphic surfaces are indistinguishable.

Remark 2.2 *As a corollary of the proposition we have that if $X : U \to S$ is a local parametrization then $X^{-1} : X(U) \to U \subset \mathbb{R}^2$ is a differentiable map. Hence S is locally diffeomorphic to a plane.*

Let S be a regular surface, $p \in S$ and let $X : U \to S$ be a local parametrization near p. If $\alpha : I \to \mathbb{R}^3$ is a regular curve such that $\alpha(I) \subset U$ (from now on we will write $\alpha : I \to S$) and $\alpha(0) = p = X(u_0, v_0)$. Then $\beta(t) = X^{-1} \circ \alpha(t)$ is a regular curve in U, that can be described as $\beta(t) = (u(t), v(t))$. Clearly

$$\alpha(t) = X \circ \beta(t)$$

in particular

$$\begin{aligned}
\alpha'(t) &= \frac{d}{dt}(X \circ \beta(t)) = \\
&= \left(\frac{\partial x}{\partial u}\frac{du}{dt} + \frac{\partial x}{\partial v}\frac{dv}{dt}, \frac{\partial y}{\partial u}\frac{du}{dt} + \frac{\partial y}{\partial v}\frac{dv}{dt}, \frac{\partial z}{\partial u}\frac{du}{dt} + \frac{\partial z}{\partial v}\frac{dv}{dt}\right) = \\
&= \frac{du}{dt}\left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right) + \frac{dv}{dt}\left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right) = \\
&= \frac{du}{dt}\frac{\partial X}{\partial u} + \frac{dv}{dt}\frac{\partial X}{\partial v}
\end{aligned}$$

It follows that any vector tangent to a curve is S passing through $p = X(u_0, v_0)$ is a linear combination of the two vectors (that do not depend on the curve, but just on the parametrization) $\frac{\partial X}{\partial u}$ and $\frac{\partial X}{\partial v}$. If $v = a\frac{\partial X}{\partial u} + b\frac{\partial X}{\partial v}$ (with $a, b \in \mathbb{R}$) is any linear combination of the two vectors then $\alpha(t) = X(at + u_0, bt + v_0)$ is a curve in S such that $\alpha(0) = p$ and $\alpha'(0) = v$. Hence any linear combination of the two vectors is tangent to some curve through p. This justifies the following definition

Definition 2.6 *Let S be a regular surface. If $p \in S$ and $X : U \to S$ is a local parametrization near p then the tangent space at p to S is the vector space T_pS spanned by the two vectors $\frac{\partial X}{\partial u}$ and $\frac{\partial X}{\partial v}$.*

Let $X_1(u,v) : U_1 \to S$ and $X_2(u',v') : U_2 \to S$ be two parametrizations near p. Then

$$X_1(u,v) = X_2 \circ X_{12}(u,v) = X_2(u'(u,v), v'(u,v))$$

hence

$$\frac{\partial X_1}{du} = \frac{\partial X_2}{du'}\frac{\partial u'}{du} + \frac{\partial X_2}{dv'}\frac{\partial u'}{du}, \qquad \frac{\partial X_1}{dv} = \frac{\partial X_2}{du'}\frac{\partial u'}{dv} + \frac{\partial X_2}{dv'}\frac{\partial u'}{dv}$$

or

$$\begin{pmatrix} \frac{\partial X_1}{du} \\ \frac{\partial X_1}{dv} \end{pmatrix} = \begin{pmatrix} \frac{\partial u'}{du} & \frac{\partial v'}{du} \\ \frac{\partial u'}{dv} & \frac{\partial v'}{dv} \end{pmatrix} \begin{pmatrix} \frac{\partial X_2}{du'} \\ \frac{\partial X_2}{dv'} \end{pmatrix}. \tag{2.1}$$

Since X_{12} is a diffeomorphism the differential of X_{12} is a linear isomorphism. Hence $\frac{\partial X_1}{du}, \frac{\partial X_1}{dv}$ and $\frac{\partial X_2}{du'}, \frac{\partial X_2}{dv'}$ span the same linear subspace of $\mathbb{R}^3$. It follows that the definition of the tangents space at p does not depend on the choice of a parametrization near p.

Example 2.8 *Let $\pi = \{(x,y,z) \in \mathbb{R}^3 : ax + by + cz + d = 0\}$ be a plane in $\mathbb{R}^3$. Then $(a,b,c) \neq (0,0,0)$ and, assuming $c \neq 0$ we have a global parametrization $U = \mathbb{R}^2 \to \pi$ given by $X(u,v) = (u, v, -\frac{1}{c}(au + bv - d))$. Hence the tangent space at a point $X(u,v)$ is spanned by the two vectors $(1,0,-\frac{a}{c})$, $(0,1,-\frac{b}{c})$. Note that the tangent space is a plane parallel to π passing through the origin.*

Example 2.9 *Consider the local parametrization of the upper hemisphere in $S^2(1)$ given by*

$$X(u,v) = (u, v, \sqrt{1 - u^2 - v^2})$$

where $(u,v) \in U = \{(u,v) \in \mathbb{R}^2 : u^2 + v^2 < 1\}$. Then

$$\frac{\partial X}{du} = (1, 0, -\frac{u}{\sqrt{1-u^2-v^2}}), \qquad \frac{\partial X}{du} = (0, 1, -\frac{v}{\sqrt{1-u^2-v^2}}).$$

This implies, in particular

$$X(u,v) \cdot \frac{\partial X}{du} = X(u,v) \cdot \frac{\partial X}{dv} = 0$$

i.e. the two vectors that span the tangent space at a point p are orthogonal to p itself. This is true for every point of the sphere. In other words $T_pS^2(1) = p^\perp$.

Example 2.10 *Consider the parametrization of a rotational surface given by*

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

where the generating curve is defined in an interval I *and* $(u,v) \in U = \{(u,v) \in \mathbb{R}^2 : u \in (0,2\pi), v \in I\}$. *Then*

$$\frac{\partial X}{du} = (-f(v)\sin(u), f(v)\cos(u), 0),$$

$$\frac{\partial X}{du} = (f'(v)\cos(u), f'(v)\sin(u), g'(v)).$$

form a basis of the tangent space to the surface.

Note that, if we have a local parametrization X near $p = X(u_0, v_0)$ then we may consider the curves

$$\alpha_1(t) = X(u_0 + t, v_0), \qquad \alpha_2(t) = X(u_0, v_0 + t)$$

it is easy to verify that the tangent vector to α_1 is always given by $\frac{\partial X}{du}(\alpha_1(t))$ and the tangent vector to α_2 is $\frac{\partial X}{dv}(\alpha_2(t))$. These curves are called coordinate curves in S. In particular

$$\frac{\partial X}{du}(p) = dX(u_0, v_0)(1,0), \qquad \frac{\partial X}{dv}(p) = dX(u_0, v_0)(0,1)$$

and the tangent space to S at p is the image of $\mathbb{R}^2 = T_{(u_0,v_0)}\mathbb{R}^2$ under the linear map $dX(u_0, v_0)$. Thus $dX(u_0, v_0)$ is linear isomorphism between $\mathbb{R}^2$ and the tangent space at p to S. The linear map $dX^{-1}(p) : T_pS \to \mathbb{R}^2$ is then well defined.

Let $f : S_1 \to S_2$ be a smooth map between surfaces. Then we want to define the differential of f at a point p. This will be a linear map between T_pS_1 and $T_{f(p)}S_2$. Let $\alpha : I \to S_1$ be a curve such that $\alpha(t_0) = p$ and $\alpha'(t_0) = v \in T_pS_1$. We know that we can describe it locally as $\alpha(t) = X(u(t), v(t))$. Then $\beta = f \circ \alpha$ is a curve on S_2 such that $\beta(t_0) = f(p)$ we define

$$df(p)(v) = \beta'(t_0)$$

it is possible to prove that $df(p)$ is well defined, i.e. the image of v does not depend on the choice of the curve α and this is a linear map $T_pS_1 \to T_{f(p)}S_2$. In fact we can write

$$df(p) = (d(f \circ X))(u_0, v_0) \circ (dX)^{-1}(p)$$

which is a composition of well defined linear maps.

This is a good definition as many properties that are known to be true in the case of maps $\mathbb{R}^n \to \mathbb{R}^m$ remain true. In particular:

Proposition 2.4 *Let $f : S_1 \to S_2$, $g : S_2 \to S_3$ be smooth map between surfaces. Then, if $p \in S_1$*

(i) $d(g \circ f)(p) = dg(f(p)) \circ df(p)$
(ii) *If $df(p)$ is a linear isomorphism then f is a local diffeomorphism of a neighborhood of p in S onto a neighborhood of $f(p)$.*
(iii) *If S is connected and $df(p) = 0$ for all p in S then f is constant.*
(iv) *If $f : S \to \mathbb{R}$ has a local maximum or minimum at p then $df(p) = 0$.*

We will not prove these facts here, but using them we can easily prove one a fundamental result in linear algebra, at least in the case of $\mathbb{R}^3$, that we will use later in these notes:

Theorem 2.3 *Let $v \to A \cdot v$ be a linear map in $\mathbb{R}^3$. If the matrix A is symmetric then there is an orthonormal basis of $\mathbb{R}^3$, $\{e_1, e_2, e_3\}$ such that $A \cdot e_i = \lambda_i e_i$, i.e. e_i are eigenvectors for the linear map with eigenvalues λ_i.*

Proof: Define a map $f : S^2(1) \to \mathbb{R}$ by $f(v) = (A \cdot v) \cdot v$ (i.e. we take the dot product of $A \cdot v$ and v). This is a smooth map and if v is a point such that $df(v) = 0$, we have that, for every curve α on $S^2(1)$ such that $\alpha(0) = v$

$$\begin{aligned} 0 &= df(v)(\alpha'(0)) = \frac{d}{dt} f(\alpha(t))_{t=0} = \frac{d}{dt}((A \cdot \alpha(t)) \cdot \alpha(t)) = \\ &= (A \cdot \alpha'(0)) \cdot v + (A \cdot v) \cdot \alpha'(0) \end{aligned}$$

the fact that A is symmetric implies that, for every $v, w \in \mathbb{R}^3$,

$$(A \cdot v) \cdot w = (A \cdot v)^t \cdot w = v^t \cdot A^t \cdot w = v^t \cdot A \cdot w = v \cdot (A \cdot w)$$

hence $df(v) = 0$ if and only if, for every vector w, tangent to $S^2(1)$ in v, we have $(A \cdot v) \cdot w = 0$, i.e. $A \cdot v$ is orthogonal to the tangent space at v to $S^2(1)$. From the description of the tangent space to a sphere (Example 2.9), this implies that $A \cdot v$ is parallel to v, i.e. $A \cdot v = \lambda v$ for some λ (i.e. v is an eigenvector) and $f(v) = \lambda v \cdot v = \lambda$ is the corresponding eigenvalue.

If f is constant then $df(v) = 0$ for every $v \in S^2(1)$. Hence every vector is an eigenvector corresponding to the eigenvalue $f(v)$ and the proof is trivial since the linear map is just a multiple of the identity. If f is not constant, using the fact that the sphere $S^2(1)$ is compact, we have that f has a maximum at some point e_1 and a minimum at e_2 with $e_1 \neq e_2$. At these points we have $df = 0$ hence e_1 and e_2 are eigenvectors, moreover

$$f(e_1)e_1 \cdot e_2 = (A \cdot e_1) \cdot e_2 = e_1 \cdot (A \cdot e_2) = f(e_2)e_1 \cdot e_2$$

and, since f is not constant, this is possible only if $e_1 \cdot e_2$ are orthogonal. Now let $e_3 = e_1 \times e_2$. Then $\{e_1, e_2, e_3\}$ is an orthonormal basis of $\mathbb{R}^3$ and

$$(A \cdot e_3) \cdot e_1 = e_3 \cdot (A \cdot e_1) = f(e_1)e_3 \cdot e_1 = 0$$

$$(A \cdot e_3) \cdot e_2 = e_3 \cdot (A \cdot e_2) = f(e_2)e_3 \cdot e_2 = 0$$

hence $A \cdot e_3$ is orthogonal to both e_1 and e_2, so it must be parallel to e_3. Then e_3 is an eigenvector of our linear map. □

Let S be a surface. Given a basis of the tangent space T_pS at $p \in S$, we would like to add a third vector to obtain a basis of $\mathbb{R}^3$, this basis should vary 'smoothly' when we move the point p on the surface

Definition 2.7 *Let S be a surface. A vector field V on S is a smooth map $S \to \mathbb{R}^3$. If $V(p) \in T_pS$ for every p in S then we say that V is a tangent vector field. If $V(p) \perp T_pS$ we say that V is a normal vector field.*

Note that if a surface is orientable and the orientation is defined by a vector field N, then we have an opposite orientation defined by the field $-N$.

Definition 2.8 *Let S be a surface. We say that S is orientable if there exists a unit normal vector field N on S.*

Example 2.11 *Any plane and the sphere $S^2(1)$ are orientable surfaces. In the case of the plane the normal field is constant, while in the case of the sphere, we can define $N(p) = p$.*

Example 2.12 *Let $S = \{(f(v)\cos(u), f(v)\sin(u), g(v)), u \in [0, 2\pi], v \in I\}$ be a rotational surface. Then, taking the cross product of the two vectors that span the tangent space (see Example 2.10) we obtain a vector that is orthogonal to the tangent space:*

$$V = (f(v)g'(v)\cos(u), f(v)g'(v)\sin(u), -f'(v))$$

then $||V||^2 = f(v)^2(g'(v)^2 + f'(v)^2) \neq 0$ since we assume that the surface is regular and $f(v) > 0$. Hence we can normalize and define an orientation of S by choosing $N = V/||V||$.

Example 2.13 *Let $S = \{(\cos(u), \sin(u), v), u \in [0, 2\pi], v \in \mathbb{R}\}$ be a cylinder. Then, as one may deduce by the formula in the previous example, or by direct computation, the normal field at a point p is just the projection of p onto the xy plane: $N = (\cos(u), \sin(u), 0)$.*

Example 2.14 *It is possible to prove that the Moebius strip, defined by*

$$S = ((1 + \frac{v}{2}\cos(\frac{u}{2}))\cos(u), (1 + \frac{v}{2}\cos(\frac{u}{2}))\sin(u), fracv2\sin(\frac{u}{2}))$$

is a non-orientable surface: the intersection of S with the xy plane is a circle $\alpha(t)$ and, assuming that S is orientable, we would have a smooth normal field $N(\alpha(t))$ at the points of the circle. However it is possible to show that if we start with $N(\alpha(0))$ and we move the point around the circle once, when we are back to $\alpha(2\pi) = \alpha(0)$ then $N(2\pi) = -N(0)$ hence N is not even continuous. This is not a proof and we will not give a rigorous proof here.

Note that orientability is a global property for a surface. Given a local parametrization $X : U \to S$ it is always possibile to define a smooth normal vector at the points of $X(U)$, as we have done in the case of rotational surfaces, by letting

$$N(X(u,v)) = \frac{1}{\left|\left|\frac{\partial X}{\partial u} \times \frac{\partial X}{\partial v}\right|\right|}\frac{\partial X}{\partial u} \times \frac{\partial X}{\partial v}. \tag{2.2}$$

Hence any surface that can be covered by just one local parametrization (for example the graphs) is orientable. Problems may arise when we try to 'glue' together the normal fields defined in different parametrizations.

(i) Let S be orientable with the orientation is defined by a normal field $\tilde{N}$. If $X : U \to S$ is a local parametrization then we have a normal field N on $X(U)$ defined by (2.2). We may have $N = \pm\tilde{N}$. If the plus sign occurs then we say that the parametrization is compatible with the orientation. In the other case we may consider a new parametrization $Y : U' \to S$ defined by $Y(u,v) = X(v,u)$ (i.e. we switch the role of u and v). Then the cross product in (2.2) will change sign and the parametrization will be compatible with the orientation.

(ii) Given two different parametrizations X_1 and X_2 near a point $p \in S$, we have two normal fields N_1 and N_2 defined by the formula (2.2). Clearly, at a point $p \in S$ we have $N_1(p) = \pm N_2(p)$. If $N_1(p) = N_2(p)$ for all p we say that the two charts are define the same orientation on S. In the case of the Moebius strip it is possible to cover the surface with two local parametrizations, but the intersection of the images of X_1 and X_2 is not connected, in one of the two connected components the orientation is the same, on the other one we have opposite orientations. It is possible to prove that a surface S that can be covered by local parametrization $X_i(U_i)$ such that $X_i(U_i) \cap X_j(U_j)$ is connected is orientable.

(iii) Given two different parametrizations X_1 and X_2 near a point $p \in S$ we have a diffeomorphism X_{12} such that $X_1 = X_2 \circ X_{12}$ (where this map is well defined). Then an easy computation following from (2.1), shows that the normal vectors defined by (2.2) coincide if and only if

$$\det \begin{pmatrix} \frac{\partial u'}{du} & \frac{\partial v'}{du} \\ \frac{\partial u'}{dv} & \frac{\partial v'}{dv} \end{pmatrix} > 0.$$

This gives us a way of checking if two local parametrizations are compatible without computing the normal vector. In particular a surface S is orientable if and only if S can be covered by local parametrizations such that the determinant of the change of coordinates is positive.

We conclude this section with the following important fact

Theorem 2.4 *Every compact surface $S \subset \mathbb{R}^3$ is orientable.*

Note that there are smooth two dimensional manifolds that are non-orientable, like the projective space or the Klein bottle. But these manifolds cannot be smoothly embedded in $\mathbb{R}^3$.

The first fundamental form

Given a symmetric bilinear form F on a vector space V, we have a quadratic form, that we will denote again by $F : V \to \mathbb{R}$, defined by $F(v) = F(v, v)$. Knowing $F(v)$ for each $v \in V$ we can recover $F(v, w)$ for $v, w \in V$ in fact from

$$F(v+w) = F(v+w, v+w) = F(v, v)+2F(v, w)+F(w, w) = F(v)+2F(v, w)+F(w)$$

we have the polarization formula

$$F(v, w) = \frac{1}{2}(F(v + w) - F(v) - F(w).$$

For example, if $F(v, w)$ is the standard dot product in $\mathbb{R}^n$ then $F(v)$ is the squared norm of v. If $\{v_1, \ldots, v_n\}$ is a basis of V then the form is completely determined by the matrix

$$F = \begin{pmatrix} F(v_1, v_1) & F(v_1, v_2) & \ldots & F(v_1, v_n) \\ F(v_2, v_1) & F(v_2, v_2) & \ldots & F(v_2, v_n) \\ \vdots & & \ddots & \\ F(v_n, v_1) & & & F(v_n, v_n) \end{pmatrix}$$

In fact, if $v = a_1 v_1 + \ldots + a_n v_n$ and $w = b_1 v_1 + \ldots + b_n v_n$ then

$$F(v, w) = \begin{pmatrix} a_1 & & a_n \end{pmatrix} \cdot F \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

We introduce a simplified notation for the functions defined on S and their derivatives. If $F : S \to \mathbb{R}^n$ is a smooth function then, instead of writing $F(p)$ for $p \in S$, we will use a local parametrization $X : U \to S$ such that $p \in X(U)$ to write $F(p) = F(u,v)$, i.e. we identify F with $F \circ X$. We will also replace the standard notation for the partial derivatives if F is a function of u and v we let

$$F_u = \frac{\partial F}{\partial u}, \qquad F_v = \frac{\partial F}{\partial v}.$$

In particular, given a local parametrization $X : U \to S$, X_u and X_v form a basis of the tangent space at the points of $X(U)$.

We now want to define two bilinear symmetric (or, equivalently, quadratic) forms on the tangent space of a regular surface.

Definition 2.9 *Let S be a regular surface. Then the first fundamental form of S at p is the bilinear symmetric form defined on T_pS by the restriction of the standard euclidean product in $\mathbb{R}^3$:*

$$I_p(v,w) = v \cdot w$$

for every $v, w \in T_pS$.

Let $p \in S$ and let $X : U \to S$ be a local parametrization of S such that $p = X(u_0, v_0)$. Then $X_u(u_0,v_0)$ and $X_v(u_0,v_0)$ form a basis of T_pS, and the first fundamental form at p is completely determined by the functions

$$E = X_u \cdot X_u, \qquad F = X_u \cdot X_v, \qquad G = X_v \cdot X_v.$$

In fact, if $w_1 = a_1X_u + a_2X_v, w_2 = b_1X_u + b_2X_v, \in T_pS$, we have

$$I_p(w_1,w_2) = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \cdot \begin{pmatrix} E & F \\ F & G \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Note that, since the Euclidean scalar product is positive definite, the matrix representing the second fundamental form is positive definite, i.e. $E > 0, G > 0, EG - F^2 > 0$. The matrix associated to the first fundamental form depend on the parametrization. If $X_1 : U_1 \to S$ (with coordinates (u,v)) and $X_2 : U_2 \to S$ (with coordinates (u',v')) are two local parametrization in a neighborhood of a point $p \in S$ then we have a smooth change of coordinates $X_{12}(u,v) = (\phi_1(u,v), \phi_2(u,v)$ and, from $X_1 = X_2 \circ X_{12}$ follows

$$X_{1u} = X_{2u'}\phi_{1u} + X_{2v'}\phi_{2u}, \qquad X_{1v} = X_{2u'}\phi_{1v} + X_{2v'}\phi_{2v}$$

If we compute the coefficient of the first fundamental form we get

$$\begin{aligned} E &= X_{1u} \cdot X_{1u} = (X_{2u'}\phi_{1u} + X_{2v'}\phi_{2u}) \cdot (X_{2u'}\phi_{1u} + X_{2v'}\phi_{2u}) = \\ &= \phi_{1u}^2 E' + 2\phi_{1u}\phi_{2u}F' + \phi_{2u}^2 G' \end{aligned}$$

$$\begin{aligned} F &= X_{1u} \cdot X_{1v} = (X_{2u'}\phi_{1u} + X_{2v'}\phi_{2u}) \cdot (X_{2u'}\phi_{1v} + X_{2v'}\phi_{2v}) = \\ &= \phi_{1u}\phi_{1v}E' + (\phi_{1u}\phi_{2v} + \phi_{2u}\phi_{1v})F' + \phi_{2u}\phi_{2v}G' \end{aligned}$$

$$\begin{aligned} G &= X_{1v} \cdot X_{1v} = (X_{2u'}\phi_{1v} + X_{2v'}\phi_{2v}) \cdot (X_{2u'}\phi_{1v} + X_{2v'}\phi_{2v}) = \\ &= \phi_{1v}^2 E' + 2\phi_{1v}\phi_{2v}F' + \phi_{2v}^2 G' \end{aligned}$$

or, in matrix form

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix} = \begin{pmatrix} \phi_{1u}) & \phi_{2u}) \\ \phi_{1v}) & \phi_{2v}) \end{pmatrix} \cdot \begin{pmatrix} E' & F' \\ F' & G' \end{pmatrix} \cdot \begin{pmatrix} \phi_{1u}) & \phi_{1v}) \\ \phi_{2u}) & \phi_{2v}) \end{pmatrix}$$

or

$$I_1(p) = dX_{12}^t \cdot I_2 \cdot dX_{12}.$$

This formula will play a role later, when we will discuss integration on surfaces. Now we want to show that the first fundamental form is related to the measure of distances on a surface. We start from the length of a curve: let $\alpha : [a, b] \to S$ be a regular curve. Then we know that we may write $\alpha(t) = X(u(t), v(t))$ hence

$$||\alpha'||^2 = ||u'X_u + v'X_v||^2 = I(u'X_u + v'X_v|) = u'^2E + 2u'v'F + v'^2G$$

and

$$L(\alpha) = \int_a^b \sqrt{u'^2E + 2u'v'F + v'^2G}.$$

Since the distance between two points on S is the infimum of the lengths of the curves joining the two points, it is clear that the first fundamental curve completely determines the distance between points on S. A f map between two surfaces in $\mathbb{R}^3$ is an isometry if the distance between points is preserved by f. We have a complete description of the isometries of $\mathbb{R}^3$. In particular all these isometries are smooth maps (we have only proved that isometries are continuous) and they preserve the length of curves. This is true in a much more general context, we state the result for the case of surfaces:

Theorem 2.5 *[Myers-Steenrod] A map f between two surfaces is an isometry if and only if f is a smooth map that preserves the length of curves.*

We will need an infinitesimal version of this result

Lemma 2.1 *A differentiable map $f : S_1 \to S_2$ between two surfaces is an isometry if and only if $df : T_pS_1 \to T_{f(p)}S_2$ preserves the scalar product.*

Proof: If f is an isometry, $p \in S_1$ and $v \in T_pS_1$, let $\alpha : I \to S_1$ be a regular curve such that $\alpha(t_0) = p$, $\alpha'(t_0) = v$. Then, for t close to t_0

$$\int_{t_0}^{t} ||\alpha'(s)||\, ds = L(\alpha)_{t_0}^{t} = L(f \circ \alpha)_{t_0}^{t} = \int_{t_0}^{t} ||df(p)(\alpha'(s))||\, ds.$$

If we take the derivative w.r to t, at $t = t_0$, we get

$$||v|| = ||\alpha'(t_0)|| = ||df(p)(\alpha'(t_0))|| = ||df(p)(v)||$$

hence $df(p)$ preserves the norm of the vectors. Using the polarization formula we can conclude that $df(p)$ preserves the scalar product.

For the opposite implication, let $\beta = f \circ \alpha$, then we have

$$L(\beta) = \int_I ||\beta'(t)||\, dt = \int_I ||df(\alpha'(t))||\, dt = \int_I ||\alpha'(t)||\, dt = L(\alpha)$$

and f is an isometry since it preserves the length of the curves. □

An immediate consequence of this infinitesimal description of the isometries between the surfaces is a relation between the two first fundamental forms. Here we only assume that f is a local isometry, i.e. for every $p \in S_1$ there is a neighborhood V of p in S_1 such that f is an isometry between V and $f(V)$:

Corollary 2.1 *A differentiable map $f : S_1 \to S_2$ between two surfaces is a local isometry if and only if*

$$I_p(w_1, w_2) = I_{f(p)}(df(p)(w_1), df(p)(w_2))$$

holds for every $p \in S_1$ and $w_1, w_2 \in T_pS_1$.

This allows us to construct maps that are local isometries between surfaces but are not induced by isometries of $\mathbb{R}^3$, i.e. we have intrinsic isometries between surfaces:

Corollary 2.2 *Let S_1 and S_2 be regular surfaces parametrized by $X_1 : U \to S_1$ and $X_2 : U \to S_2$. If*

$$E_1 = E_2, \quad F_1 = F_2, \quad G_1 = G_2$$

then $\circ X_1 X_2^{-1}$ is an isometry between S_1 and S_2.

Proof: Let $\beta : I \to U$ be a regular curve and let $\alpha_1 = X_1 \circ \beta$, $\alpha_2 = X_2 \circ \beta$. Then $\alpha_1 = X_{12}$

$$\begin{aligned} ||\alpha_1'||^2 &= ||u'X_{1u} + v'X_{1v}||^2 = u'^2E_1 + 2u'v'F_1 + v'^2G_1 = \\ &= u'^2E_2 + 2u'v'F_2 + v'^2G_2 = ||u'X_{2u} + v'X_{2v}||^2 = ||\alpha_2'||^2. \end{aligned}$$

Then f preserves the length of tangent vectors, the length of curves is then also preserved. $\square$

Example 2.15 *Let $U = \{(u, v) \in \mathbb{R}^2 : u \in (0, 2\pi)\}$ and let*

$$X_1(u, v) = (u, v, 0), \qquad X_2(u, v) = (\cos(u), \sin(u), v)$$

Then $S_1 = X_1(U)$ is an open subset of a plane in $\mathbb{R}^3$, while $S_2 = X_2(U)$ is a subset of a cylinder. We have

$$X_{1u} = (1, 0, 0), \quad X_{1v} = (0, 1, 0), \quad X_{2u} = (-\sin(u), \cos(u), 0), \quad X_{2v} = (0, 0, 1)$$

and $E_1 = E_2 = G_1 = G_2 = 1$, $F_1 = F_2 = 0$. Hence S_1 and S_2 are isometric. Note that this isometry cannot be induced by an isometry of $\mathbb{R}^3$ since every isometry of $\mathbb{R}^3$ maps planes into planes. The plane and the cylinder are not globally isometric, in fact the two surfaces are topologically distinct, the first fundamental group of the plane is trivial while the one of the cylinder is not: every circle in $\mathbb{R}^2$ can be continuously deformed to a point while there are circles in the cylinder that are not contractible.

The first fundamental form of a surface allows to compute distances on the surface, by integrating the norm of tangent vectors to curves. Hence it is related to the 'intrinsic' geometry of the surface, i.e. to quantities that can be computed without knowing how the surface is embedded in $\mathbb{R}^3$.

The second fundamental form

We studied the curves in $\mathbb{R}^3$ using the Frenet frame. The starting point is the variation of the tangent space to the curve. We now want to study the variation of the tangent space to a two dimensional surface S. Since any surface is locally orientable and the tangent space is completely determined by the normal vector, it is enough to study the variation of a unit normal field on S. This will lead us to the definition of another bilinear form on T_pS.

Definition 2.10 *Let S be an orientable surface. The Gauss map on S is the map $N : S \to S^2(1)$ defined by $p \to N(p)$.*

Example 2.16 *If S is a plane then the Gauss map is constant. On $S^2(1)$ the Gauss map is the identity. On a cylinder the Gauss map is not constant and the image of the Gauss map is a circle (see Example 2.13).*

The Gauss map is a smooth map and we may consider its differential at a point $p \in S$. This will be a linear map $dN : T_pS \to T_{N(p)}S^2(1)$. But the characterization of the tangent space to $S^2(1)$ (Example 2.9) tells us that $T_{N(p)}S^2(1)$ is orthogonal to $N(p)$. The space T_pS is orthogonal to $N(p)$ by definition, hence the two spaces coincide.

Definition 2.11 *Let S be an orientable surface. The Weingarten map on S at the point p is the linear map $S_p : T_pS \to T_pS$ defined by*

$$S_p = -dN(p).$$

The Gauss curvature $K(p)$ and the mean curvature $H(p)$ of S at p are defined by

$$K(p) = \det(S_p), \qquad H(p) = \frac{1}{2}\operatorname{tr}(S_p).$$

Note that given an endomorphism, the determinant and the trace can be computed using any matrix associated to the linear map w.r. to a basis of the vector space. Then we define the second fundamental form of S by

$$II_p(w_1, w_2) = S_p(w_1)\cdot w_2 = -dN(p)(w_1)\cdot w_2.$$

With respect to the $\{X_u, X_v\}$ basis we have

$$d(N)(X_u) = \frac{d}{dt}N(u_0+t, v_0) = N_u, \qquad d(N)(X_v) = \frac{d}{dt}N(u_0, v_0+t) = N_v$$

and

$$d(N)(X_u)\cdot X_v = N_u\cdot X_v = \frac{d}{du}(N\cdot X_v) - N\cdot X_{vu} = -N\cdot X_{vu}$$

$$d(N)(X_v)\cdot X_u = N_v\cdot X_u = \frac{d}{dv}(N\cdot X_u) - N\cdot X_{uv} = -N\cdot X_{uv} = -N\cdot X_{uv}$$

if we let

$$e = -N_u\cdot X_u, \qquad f = -N_u\cdot X_v = -N_v\cdot X_u, \qquad g = -N_v\cdot X_v$$

then, if $w_1 = a_1X_u + a_2X_v$ and $w_2 = b_1X_u + b_2X_v$, we have

$$II_p(w_1, w_2) = \begin{pmatrix} a_1 & a_2 \end{pmatrix}\cdot\begin{pmatrix} e & f \\ f & g \end{pmatrix}\cdot\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

In particular, the second fundamental form is a symmetric bilinear form on T_pS. Note that we can also compute the coefficients of the second fundamental form as

$$e = N\cdot X_{uu}, \qquad f = N\cdot X_{uv}, \qquad g = N\cdot X_{vv}$$

and these expressions are easier to compute since, in general, the expression for the normal vector N is complicated and here we avoid taking derivatives of this term. We now want to derive a formula for the Weingarten operator, since the

Gauss curvature and the mean curvature are defined in terms of this operator. Since $d(N)(X_u) \cdot X_v = dN(X_v) \cdot X_u$, with respect to the $\{X_u, X_v\}$ basis this linear operator is represented by a symmetric matrix

$$S_p = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$

where

$$dN(X_u) = N_u = -(a_{11}X_u + a_{21}X_v), \qquad dN(X_v) = N_v = -(a_{12}X_u + a_{22}X_v)$$

from the first equation, taking the scalar product with X_u on both sides, we get

$$e = -N_u \cdot X_u = a_{11}X_u \cdot X_u + a_{12}X_v \cdot X_u = a_{11}E + a_{12}F$$

similarly

$$f = -N_u \cdot X_v = a_{11}X_u \cdot X_v + a_{12}X_v \cdot X_v = a_{11}F + a_{12}G$$

$$g = -N_v \cdot X_v = a_{12}X_u \cdot X_v + a_{22}X_v \cdot X_v = a_{12}F + a_{22}G$$

or, in matrix form:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \cdot \begin{pmatrix} e & f \\ f & g \end{pmatrix}.$$

By taking the determinant of the right end side we obtain a formula for the Gauss curvature in terms of the coefficients of the first and the second fundamental form:

$$K = \frac{eg - f^2}{EG - F^2} \tag{2.3}$$

since

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

we can compute the product and obtain the following formula for the mean curvature:

$$H = \frac{eG - 2fF + Eg}{2(EG - F^2}.$$

Example 2.17 *Let S be a plane in $\mathbb{R}^3$. Then, at every point of $p \in S$, we have a constant normal field $N(p) = N_0$. It follows that $dN = 0$ and both the Gauss and the mean curvature vanish identically.*

Example 2.18 *Let $S = S^2(r)$ be the sphere of radius r centered at the origin. Then, for $p \in S$, we have $N(p) = \frac{1}{r}p$. It follows that $N = \frac{1}{r}I$ is a linear map, hence $dN = N$. The determinant of $-dN$, i.e. Gauss curvature, is constant equal to $\frac{1}{r^2}$, while the mean curvature is equal to $-\frac{1}{r}$.*

Example 2.19 *We have already seen (see Example 2.13) that the normal vector at $p = X(u,v) = (\cos(u), \sin(u), v)$ in the cylinder is $N(u,v) = (\cos(u), \sin(u), 0)$, we have*

$$X_u = (-\sin(u), \cos(u), 0), \qquad X_v = (0,0,1)$$

$$X_{uu} = (-\cos(u), -\sin(u), 0), \qquad X_{uv} = (0,0,0). \qquad X_{vv} = (0,0,0)$$

hence $E = G = 1$, $F = 0$, $e = -1$, $f = g = 0$ hence $K = 0$ and $H = \frac{1}{2}$.

Example 2.20 *Consider the helicoid, defined by $S = \{v\cos(u), v\sin(u), u)$ for $u \in [0, 2\pi]$ and $v \in \mathbb{R}$. We have*

$$X_u = (-v\sin(u), v\cos(u), 1), \qquad X_v = (\cos(u), \sin(u), 0)$$

and a normal field defined by

$$N = \frac{1}{\sqrt{1+v2}}(-\sin(u), \cos(u), -v).$$

We also have

$$X_{uu} = (-v\cos(u), -v\sin(u), 0), \quad X_{uv} = (-\sin(u), \cos(u), 0), \quad X_{vv} = (0,0,0)$$

hence

$$E = 1 + v^2, \qquad F = 0, \qquad G = 1$$

$$e = 0, \qquad f = \frac{1}{\sqrt{1+v^2}}, \qquad g = 0.$$

It follows that the Gauss curvature is equal to

$$K = -\frac{1}{(1+v^2)^2}$$

while the mean curvature H is zero. A surface with zero mean curved is called minimal.

We proved that the matrix associated to the Weingarten operator with respect to the basis $\{X_u, X_v\}$ is symmetric. It follows that the spectral theorem applies and it is possible to determine an orthonormal basis $\{e_1, e_2\}$ of T_pS at every point $p \in S$ such that $S_p(e_1) = k_1 e_1$, $S_p(e_2) = k_2 e_2$. The vectors e_1, e_2 are called principal directions at p while the eigenvalues k_1 and k_2 are the principal curvatures at p .

As in the case of the curves it is possible to derive geometric information on S from the knowledge of the curvature, there are cases when this is particularly efficient:

Definition 2.12 *Let S be a surface. A point $p \in S$ is umbilical if the principal curvatures are equal at p. If every point of S is umbilical then S is called umbilical.*

Theorem 2.6 *Let S be a connected umbilical surface. Then S is an open subset of a sphere or of a plane.*

Proof: Let $p \in S$. Since the eigenvalues k_1, k_2 of S_p are equal we have that, w.r. to the basis e_1, e_2, S_p is associated to the matrix

$$\begin{pmatrix} k_1 & 0 \\ 0 & k_1 \end{pmatrix} = k_1 I$$

i.e. S_p and $dN(p)$ are multiples of the identity on T_pS and we can write $dN(p) = \lambda(p)I$, where $\lambda(p)$ is a smooth function on S (in fact $\lambda(p) = -k_1(p)$). We first work in a local parametrization $X : U \to S$ of S. Let $v = aX_u + bX_v \in T_pS$ then $dN(p)(v) = \lambda v$ hence

$$\lambda(aX_u + bX_v) = dN(p)(aX_u + bX_v) = aN_u + bN_v$$

it follows that

$$N_u = \lambda X_u, \qquad N_v = \lambda X_v \tag{2.4}$$

we take the derivative of the first equation with respect to v and the derivative of the second one with respect to u:

$$N_{uv} = \lambda_v X_u + \lambda X_{uv}, \qquad N_{vu} = \lambda_u X_v + \lambda X_{vu}$$

but $N_{uv} = N_{vu}$ and $X_{uv} = X_{vu}$ hence taking the difference of the two equations we get

$$\lambda_v X_u = \lambda_u X_v$$

since X_u and X_v are linearly independent this is possible only if $\lambda_u = \lambda_v = 0$ i.e. λ is a constant function on the local parametrization of S.

We have two cases

(i) If $\lambda = 0$ then $dN = 0$ hence $N(p) = N_0$ is constant. Let $(u_0, v_0) \in U$ and let $f(u,v) = (X(u,v) - X(u_0,v_0)) \cdot N_0$. Then

$$\frac{\partial f}{\partial u} = X_u \cdot N = 0, \qquad \frac{\partial f}{\partial v} = X_v \cdot N = 0$$

hence $f(u,v)$ is constant. Since $f(u_0,v_0) = 0$ we have that $X(u,v)$ belongs to a plane orthogonal to N_0 and passing through $X(u_0,v_0)$.

(ii) If $\lambda \neq 0$ consider the function $f(u,v) = X(u,v) - \frac{1}{\lambda}N_0$. Then, using 2.4

$$\frac{\partial f}{\partial u} = X_u - \frac{1}{\lambda}N_u = 0, \qquad \frac{\partial f}{\partial v} = X_v = \frac{1}{\lambda}N_v = 0$$

and $f(u,v) = q$ is constant. Then

$$||X(u,v) = q|| = \left|\left|\frac{1}{\lambda}N(u,v)\right|\right| = \frac{1}{\lambda}$$

i.e. the points $X(u,v)$ lie on a sphere of radius $\frac{1}{\lambda}$ centered in q.

Now we want to prove that this result is global. Given two points on the surface we can connect them with a curve $[a,b] \to S$. The trace of the curve is then compact (since the image of a compact set via a continuous map is compact) and can be covered with a finite number of local parametrizations. When two parametrizations overlap we get that they must cover a portion of the same plane or of the same sphere. Hence S is a subset of a plane or of a sphere. $\square$

Now we describe the second fundamental form in different ways that will help us to derive geometric consequences from the value of the principal curvatures of a surface S. Let $\alpha : I \to S$ be a bi-regular curve parametrized by arclenght. Then $\alpha''(t) = k(t)\,\underline{n}(t)$, where $k(t)$ is the curvature of α and $\underline{n}$ is the normal vector in the Frenet frame of α.

Definition 2.13 *Let S be a regular orientable surface and let $\alpha : I \to S$ be a bi-regular curve. Then we define the normal curvature*
Curve!Normal curvature of α at a point $p = \alpha(t_0) \in S$ by

$$k_n(t_0) = k(t_0)\,\cos(\theta(t_0))$$

where k is the curvature of α and θ is the angle between the normal vector $\underline{n}$ of α and the unit normal $N(p)$ of S at p.

Note that the sign of the normal curvature depends on the choice of the orientation of S. Since $\underline{n}$ and N are both unit vectors we have an equivalent definition

$$k_n(t_0) = \alpha''(t_0) \cdot N(\alpha(t_0))$$

i.e., up to sign, k_n is the length of the projection of α'' on the normal to the surface. Denote by $N(t)$ the restriction of N to the points of $\alpha(t)$. Then, from the definition of differential if follows that $dN(\alpha'(t_0)) = \frac{d}{dt}N(\alpha(t))_{t=t_0} = N'(t_0)$. Moreover, since $\alpha'(t)$ is tangent to S, we have $\alpha'(t) \cdot N(t)$ hence

$$0 = \alpha''(t) \cdot N(t) + \alpha'(t) \cdot N'(t) = \alpha''(t) \cdot N(t) + \alpha'(t) \cdot dN(\alpha'(t)).$$

Hence

$$II(\alpha'(t_0)) = -dN(\alpha'(t_0)) \cdot \alpha'(t_0) = \alpha''(t_0) \cdot N(t_0) = k_n(t_0).$$

The left hand side of this equation only depends on the value of $v = \alpha'(t_0)$, it follows that, if $\beta(t_0) = \alpha(t_0)$ and $\beta'(t_0) = v$ then α and β have the same normal curvature at $t = t_0$ and the normal curvature is in fact a function $k_n(v)$ defined on the tangent space. In particular, to compute $k_n(v)$ we can choose any curve through p that is tangent to v. Let $\alpha(t)$ be the plane curve determined by the intersection of S with the plane through p spanned by v and $N(p)$. This curve is the normal section of S at p in the direction of p, we can parametrize α with arclength and so that $v = \alpha'(t_0)$. Since α is a plane curve the normal vector to α lies on the plane and, being orthogonal to v, must coincide with $\pm N(p)$. It follows that $\underline{n}(t_0) \cdot N(p) = \pm 1$. and $k_n = \pm k$. The normal curvature of the normal section coincides, up to sign, with the curvature of α. The value of the second fundamental form at p in the direction of $v \in T_pS$ is (up to sign) the curvature of the corresponding normal section.

Let $e_1, e_2 \in T_pS$ be the principal directions of S at p and let $v \in T_pS$ be an unit vector. Then (since the basis is orthonormal) we can write $v = \cos(\theta)e_1 + \sin(\theta)e_2$ and

$$\begin{aligned} -dN(p)(v) &= S_p(\cos(\theta)e_1 + \sin(\theta)e_2) = \cos(\theta)S_p(e_1) + \sin(\theta)S_p(e_2) = \\ &= \cos(\theta)k_1e_1 + \sin(\theta)k_2e_2 \end{aligned}$$

and we have the Euler formula for the normal curvature

$$\begin{aligned} k_n(v) &= II_p(v) = -dN(p)(v) \cdot v = \\ &= (\cos(\theta)k_1e_1 + \sin(\theta)e_2) \cdot (\cos(\theta)e_1 + \sin(\theta)e_2) = \\ &= \cos(\theta)^2k_1 + \sin(\theta)^2k_2. \end{aligned}$$

Since v is uniquely determined by θ we can write $k_n(v) = k_n(\theta)$ and, taking the derivative w.r. to θ

$$k_n'(\theta) = -2\cos(\theta)\sin(\theta)k_1 + 2\sin(\theta)\cos(\theta)k_2 = 2\cos(\theta)\sin(\theta)(k_2 - k_1)$$

if $k_1 = k_2$ then k_n is constant equal to k_1. Otherwise we have critical points of k_n only if $\cos(\theta) = 0$ or $\sin(\theta) = 0$. It follows that the principal curvatures k_1

and k_2 are the minimum and the maximum value that the normal curvature can assume at the point p.

Definition 2.14 *Let S be a regular orientable surface in $\mathbb{R}^3$. A point $p \in S$ is elliptic if $K(p) > 0$, hyperbolic if $K(p) < 0$, parabolic if $K(p) = 0$ but $H(p) \neq 0$ and planar if $K(p) = H(p) = 0$.*

If a point $p \in S$ is elliptic then $K(p) = k_1(p)k_2(p) > 0$ and k_1 and k_2 have the same sign. It follows that all the normal curvatures at p have the same sign, hence (up to a change of orientation in S), we may assume that for any normal section at p, we have $\underline{n} = N$. It follows that the normal sections are all curved 'the same way' near p and they all stay on one side of the tangent space at p to S. Conversely if $K < 0$ we have normal sections at p that lie on both sides of the tangent space at p.

Example 2.21 *The origin is an elliptic (resp. hyperbolic) point for the graph of the function $f(x,y) = x^2 + y^2$ (resp $f(x,y) = x^2 - y^2$). All the points of a cylinder are parabolic while all the points of a plane are planar. Note that the normal curvature of a curve α is determined by the second derivative of α. If the second derivative of α vanishes then we may have a planar point on S even if the surface is not a plane. For example, if we consider the rotational surface $(v\cos(u), v\sin(u), v^4)$, with $u \geq 0$ and $v \in (0, 2\pi)$, the resulting surface is regular (even if the graph of the generating function intersects the z axis) and the origin is a planar point.*

We will now describe two constructions that lead to new definitions of the normal curvature. Let $f : S \to \mathbb{R}$ be a smooth function. Then, for $p \in S$ we may consider a curve $\alpha : I \to S$ such that $\alpha(0) = p$. Let $w = \alpha'(0)$. We define the Hessian of f at p by

$$H_f(p)(w) = \frac{d^2}{dt^2}(f \circ \alpha)_{t=0}.$$

This is a map defined on T_pS. In local coordinates, if $\alpha(t) = \alpha(u(t), v(t))$, we have

$$\begin{aligned}
H_f(p)(w) &= \frac{d^2}{dt^2} f(u(t), v(t))_{t=0} = \frac{d}{dt}(\frac{\partial f}{\partial u}u' + \frac{\partial f}{\partial v}v' = \\
&= \frac{\partial^2 f}{\partial u^2}u'^2 + 2\frac{\partial^2 f}{\partial u \partial v}u'v' + \frac{\partial^2 f}{\partial v^2}v'^2 + \frac{\partial f}{\partial u}u'' + \frac{\partial f}{\partial v}v'' = \\
&= \begin{pmatrix} u' & v' \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial^2 f}{\partial u^2} & \frac{\partial^2 f}{\partial u \partial v} \\ \frac{\partial^2 f}{\partial u \partial v} & \frac{\partial^2 f}{\partial v^2} \end{pmatrix} \cdot \begin{pmatrix} u' \\ v' \end{pmatrix} + \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix} \cdot \begin{pmatrix} u'' \\ v'' \end{pmatrix} =
\end{aligned}$$

$$= \quad w^t \cdot \begin{pmatrix} \frac{\partial^2 f}{\partial u^2} & \frac{\partial^2 f}{\partial u \partial v} \\ \frac{\partial^2 f}{\partial u \partial v} & \frac{\partial^2 f}{\partial v^2} \end{pmatrix} \cdot w + \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix} \cdot \begin{pmatrix} u'' \\ v'' \end{pmatrix}.$$

The second summand depends on the choice of the curve, not only on the tangent vector w so, in general, $H_f(p)$ is not well defined. But if p is a critical point of f, i.e. $\frac{\partial f}{\partial u}(p) = \frac{\partial f}{\partial v}(p) = 0$ then the Hessian is well defined and coincides, formally, with the standard Hessian of a function $\mathbb{R}^2 \to \mathbb{R}^2$.

Let S be a regular orientable surface and let $p = X(u_0, v_0) \in S$ in a local parametrization of S. We define the function

$$f(u,v) = (X(u,v) - X(u_0,v_0)) \cdot N(u_0,v_0)$$

then

$$\frac{\partial f}{\partial u}(p) = X_u(p) \cdot N(p) = 0, \qquad \frac{\partial f}{\partial v} = X_v(p) \cdot N(p) = 0$$

and p is a critical point of f. The function f measures the projection of $X(u,v) - p$ on the normal vector at p, i.e. the (signed) height of $X(u,v)$ relative to the tangent plane at p. Let $w \in T_pS$ be a unit vector and let $\alpha : I \to S$ be a curve, parametrized by arclength, such that $\alpha(0) = p$ and $\alpha'(0) = w$. Then we have

$$II_p(w) = k_n(w) = \alpha'' \cdot N(p) = \frac{d^2}{dt^2}(\alpha(t) - \alpha(0)) \cdot N(p) = H_f(p)(w).$$

For a real valued function of several variables the Hessian at a critical point gives information about the nature of the critical point. E.g., if the Hessian is positive/negative definite then the point is a local minimum/maximum. If the point p is elliptic this is exactly the case since the principal curvatures have the same sign. It follows that the local picture of the surface, as graph with respect to the tangent space at p, is similar to the one of a paraboloid at the origin, i.e. the surface lies on one side of the tangent space.

Example 2.22 *We consider a torus $T^2 \subset \mathbb{R}^3$, described as rotational surface:*

$$X(u,v) = ((r\cos(v) + R)\cos(u), (r\cos(v) + R)\sin(u), r\sin(v))$$

for $u, v \in [0, 2\pi]$ and $R > r$. Then

$$X_u(u,v) = (-(r\cos(v) + R)\sin(u), (r\cos(v) + R)\cos(u), 0),$$

$$X_v(u,v) = (-r\sin(v)\cos(u), -r\sin(v)\sin(u), r\cos(v))$$

and

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

$$X_{uu} = (-(r\cos(v)+R)\cos(u), -(r\cos(v)+R)\sin(u), 0),$$

$$X_{uv} = (r\sin(v)\sin(u), -r\sin(v)\cos(u), 0),$$

$$X_{vv} = (r\cos(v)\cos(u), -r\sin(u)\cos(v), -r\sin(v)).$$

Hence

$$E = (r\cos(v)+R)^2, \quad F = 0, \quad G = r^2$$

$$e = -(r\cos(v)+R)\cos(v), \quad f = 0, \quad g = -r$$

and the Gauss curvature is given by

$$K = \frac{\cos(v)}{r(r\cos(v)+R)}.$$

The denominator is always positive, hence we have $K > 0$ at the points where $\cos(v) > 0$. This are the points on the 'exterior' of the torus, and it is clear that the surface lies on one side of the tangent space. The points with $K = 0$ are the ones of the top and lower circles. The tangent space at these points contains the circles. The points with $K < 0$ are the ones in the 'interior' of the torus. There the tangent space intersects the surface.

Calculus on surfaces

Denote by $C^\infty(\mathbb{R}^n)$ the set of smooth real valued functions on $\mathbb{R}^n$ and let $w \in \mathbb{R}^n$ be a nonzero vector. Then we can associate to each function $f \in C^\infty(\mathbb{R}^n)$ its partial derivative in the direction of w: $\frac{\partial f}{\partial w}$. This is another function in $C^\infty(\mathbb{R}^n)$ and we can think of the vector w as an operator on the set of smooth functions:

$$w(f) = \frac{\partial f}{\partial w} = df(w).$$

We are not proving it here, but the knowledge of this map uniquely determines the vector w, for example if $f(x,y) = x^2 + 3xy - y + 1$ and we know that $\frac{\partial f}{\partial w} = 2x + 3y$ then it is clear that $w = (1,0)$ and we are taking the derivative w.r. to the x variable. This operator has the following properties

(i) Linearity: $\forall f,g \in C^\infty(\mathbb{R}^n)$, $\forall a,b \in \mathbb{R}$: $w(af+bg) = aw(f) + bw(g)$.
(ii) Leibnitz rule: $\forall f,g \in C^\infty(\mathbb{R}^n)$: $w(fg) = w(f)g + fw(g)$.

We can replace the fixed vector v with a vector field X on $\mathbb{R}^n$ and let

$$Xf(p) = \frac{\partial f}{\partial X(p)}(p) = df(p)(X(p))$$

and a vector field X can be seen as a differential operator on $C^\infty(\mathbb{R}^n)$. Denote by $\mathfrak{X}(\mathbb{R}^n)$ the set of smooth vector fields on $\mathbb{R}^n$. Then we have the following properties:

(i) Linearity: $\forall X, Y \in \mathfrak{X}(\mathbb{R}^n)$, $\forall f \in C^\infty(\mathbb{R}^n)$, $\forall a, b \in \mathbb{R}$: $(aX + by)(f) = aX(f) + bY(f)$.
(ii) Tensoriality: $\forall X \in \mathfrak{X}(\mathbb{R}^n)$, $\forall f, g \in C^\infty(\mathbb{R}^n)$: $(fX)(g) = fX(g)$.

From the last property follows, taking $f(p) = 0$, that if $X(p) = 0$ then $X(g) = 0$ at p. Hence if two vector fields X_1, X_2 coincide at p then $0 = X_1(g)(p) - X_2(g)(p)$, i.e. the value of $X(g)$ at p only depends on the value of X at p.

Denote by $\{e_1, \dots, e_n\}$ the standard basis of $\mathbb{R}^n$. If

$$X(x_1, \dots, x_n) = (a_1(x_1, \dots, x_n), \dots, a_n(x_1, \dots, x_n) = \sum_i a_i e_i$$

then X acts on a function f by

$$X(f) = \sum_i a_i e_i(f) = \sum_i a_i \frac{\partial f}{\partial x_i}$$

If $Y = (b_1(x_1, \dots, x_n), \dots, b_n(x_1, \dots, x_n)$ is another vector field then we may take the derivative, with respect to X, of the function $Y(f)$:

$$XY(f) = X\left(\sum_j b_j \frac{\partial f}{\partial x_j}\right) = \sum_i a_i \left(\sum_j \frac{\partial b_j}{\partial x_i}\frac{\partial f}{\partial x_j} + b_j \frac{\partial f}{\partial x_j \partial x_i}\right)$$

similarly

$$YX(f) = Y\left(\sum_j a_j \frac{\partial f}{\partial x_j}\right) = \sum_i b_i \left(\sum_j \frac{\partial a_j}{\partial x_i}\frac{\partial f}{\partial x_j} + a_j \frac{\partial f}{\partial x_j \partial x_i}\right)$$

hence

$$(XY - YX)(f) = \sum_{ij} \left(a_i \frac{\partial b_j}{\partial x_i} - b_i \frac{\partial a_j}{\partial x_i}\right) \frac{\partial f}{\partial x_j \partial x_i}$$

and $XY - YX$ is a well defined vector field (while XY and YX are not, due to the presence of the second derivatives) denoted by $[X, Y]$ and called Lie bracket of the field X and Y. Note that if the component of a_i, b_j of the two vector fields are constant, then $[X, Y] = 0$ (hence $XY = YX$), e.g. $[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] = 0$ and this reflects the well know fact that the mixed partial derivatives with respect to coordinate fields are equal.

A vector field $Y = (b_1(x_1, \dots, x_n), \dots, b_n(x_1, \dots, x_n)$ is defined by a set of n real valued functions, so we may define the derivative of the vector field Y with respect to the vector field X by

$$D_X Y = (X(b_1), \ldots, X(b_n))$$

then $D_X Y$ is again a vector field in $\mathbb{R}^n$.

Example 2.23 *Let*

$$X = (xy, -2y) = (xy)\frac{\partial}{\partial x} - 2y\frac{\partial}{\partial y}, \qquad Y = (x, -3y^2) = x\frac{\partial}{\partial x} - 3y^2\frac{\partial}{\partial y}.$$

Then

$$\begin{aligned} D_X Y &= (X(x), X(-3y^2)) = \\ &= ((xy)\frac{\partial}{\partial x}(x) - 2y\frac{\partial}{\partial y}(x), (xy)\frac{\partial}{\partial x}(-3y^2) - 2y\frac{\partial}{\partial y}(-3y^2)) = \\ &= (xy, 12y^2) = xy\frac{\partial}{\partial x} + 12y^2\frac{\partial}{\partial y} \end{aligned}$$

for example, at the point $p = (1, 2)$*, we have* $D_X Y(p) = 2\frac{\partial}{\partial x} + 48\frac{\partial}{\partial y}$.

It is possible to give a different, more geometric, description of $D_X Y$. Let $p \in \mathbb{R}^n$ and let $v = X(p)$. If α is a regular curve in $\mathbb{R}^n$ such that $\alpha(0) = p$ and $\alpha'(0) = v$ then we may consider the restriction $Y(t)$ of the vector field Y to the points of $\alpha(t)$ and

$$D_X Y(p) = \frac{d}{dt} Y(\alpha(t))_{t=0}.$$

Example 2.24 *We consider* X, Y *from the previous example and let* $p = (1, 2)$*. Then* $X(p) = (2, -4)$ *and we may consider the curve* $\alpha(t) = (1 + 2t, 2 - 4t)$ *passing through* p *at time* $t = 0$ *and such that* $\alpha'(0) = X(p)$*. We have*

$$Y(t) = (1 + 2t, -3(2 - 4t)^2)$$

and

$$D_X Y(p) = \frac{d}{dt}(1 + 2t, -3(2 - 4t)^2)_{t=0} = (2, 48) = 2\frac{\partial}{\partial x} + 48\frac{\partial}{\partial y}.$$

In particular the value of $D_X Y$ at a point p can be determined knowing only the value of X at p and the values of Y along any curve through p tangent at $X(p)$. The following properties of the derivative of a vector field are easy to prove:

(i) Linearity: $\forall X, Y, Z \in \mathfrak{X}(\mathbb{R}^n)$, $\forall a, b \in \mathbb{R}$:

$$D_X(aY + bZ) = aD_X Y + bD_X Z, \qquad D_{(aX+bY)} Z = aD_X Z + bD_Y Z.$$

(ii) Leibnitz rule: $\forall X, Y \in \mathfrak{X}(\mathbb{R}^n)$, $\forall f \in C^\infty(\mathbb{R}^n)$:

$$D_X(fY) = X(f)Y + fD_XY.$$

(iii) Tensoriality: $\forall X, Y \in \mathfrak{X}(\mathbb{R}^n)$, $\forall f \in C^\infty(\mathbb{R}^n)$:

$$D_{(fX)}Y = fD_XY.$$

(iv) Torsion free: $\forall X, Y \in \mathfrak{X}(\mathbb{R}^n)$:

$$D_XY - D_YX = [X, Y].$$

(v) Compatibility with the metric: $\forall X, Y, Z \in \mathfrak{X}(\mathbb{R}^n)$:

$$X(Y \cdot Z) = (D_XY) \cdot Z + Y \cdot (D_XZ).$$

Since, to define D_XY, Y does not have to defined in an open set of $\mathbb{R}^n$, it make sense to study D_XY for $X, Y, \in \mathfrak{X}(S)$, where S is a regular surface in $\mathbb{R}^3$. The result is, in general, a vector in $\mathbb{R}^n$ that it is not tangent to S, while we would like to obtain an intrinsic object. A result in this direction is the following

Proposition 2.5 *Let S be a surface and let $X, Y \in \mathfrak{X}(S)$, then the bracket of the two vector fields (computed in $\mathbb{R}^3$) is tangent to S, i,e*

$$[X, Y] \in \mathfrak{X}(S).$$

If, in particular, X and Y are coordinate vector fields, then it is possible to prove that $[X, Y] = 0$.

Example 2.25 *Let $X : U \to S$ be a local parametrization of a regular surface S. Let $X = X_u$ and $Y = X_v$, then*

$$D_XY = \frac{\partial}{\partial u}\left(\frac{\partial x_1}{\partial v}, \frac{\partial x_2}{\partial v}, \frac{\partial x_3}{\partial v}\right) = X_{uv}$$

and we know that, in general, X_{uv} has a component along the normal to the surface.

Definition 2.15 *Let S be a regular surface and let $X, Y \in \mathfrak{X}(S)$ then we define the covariant derivative of Y with respect to X by*

$$\nabla_XY = (D_XY)^{tang}$$

where Z^{tang} denotes the orthogonal projection on the tangent space to S.

Again, the value at p of $\nabla_X Y$ only depends on the value of X at p and the values of Y along any curve through p tangent to $X(p)$. It is possible to verify that the 5 properties of $D_X Y$ still hold for the covariant derivative (assuming Proposition 2.5 the proof is an easy consequences of the formulas in $\mathbb{R}^n$).

It is convenient to introduce a different notation that will allow simplify some formula and the generalization to higher dimensions. Let V be an Euclidean vector space (i.e. a vector space equipped with a positive definite symmetric bilinear form), we denote the scalar product of $X, Y \in V$ by $< X, Y >$ and if $\{v_1, \ldots, v_n\}$ is a basis of V we let

$$g_{ij} =< v_i, v_j >$$

for $i, j = 1, \ldots, n$. This defines a symmetric positive definite matrix G. Then G is invertible and we denote by g^{kl} the entries of G^{-1}.

Example 2.26 *Let $v \in V$ and suppose that we know the scalar product of v with each element v_i of a basis of V. Let*

$$< v, v_i >= a_i$$

it is possibile to write v as linear combination of the elements of the basis of V: $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$. We want to determine the relations between the coefficients λ_i and a_j. We have

$$a_i =< v, v_i >=< \sum_{j=1}^{n} \lambda_j v_j, v_i >= \sum_{j=1}^{n} \lambda_j < v_j, v_i >= \sum_{j=1}^{n} g_{ij} \lambda_j$$

it follows that

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = G \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

or

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = G^{-1} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

hence

$$\lambda_j = \sum_{i=1}^{n} g^{ji} a_i.$$

This shows that it is possible to recover the coefficients of v as linear combination of the elements of a basis by knowing the scalar products of v with these elements.

We can show that the properties 1),...,5) of the covariant derivative allow us to completely determine ∇:

Proposition 2.6 *Let S be a surface and let $\tilde{\nabla}$ be an operator $\mathfrak{X}(S)\times\mathfrak{X}(S) \to \mathfrak{X}(S)$ that satisfies the properties* 1),...,5) *definded above. Then $\tilde{\nabla} = \nabla$.*

Proof: Given $X, Y \in \mathfrak{X}(S)$ it is enough to show that $\nabla_X Y$ is completely determined by the properties 1),...,5) (hence it is unique) showing that $< \nabla_X Y, Z >$ is determined for every $Z \in \mathfrak{X}(S)$. We have

$$\begin{aligned}
< \nabla_X Y, Z > &= X < Y, Z > - < Y, \nabla_X Z >= \\
&= X < Y, Z > - < Y, [X, Z] > - < Y, \nabla_Z X >= \\
&= X < Y, Z > - < Y, [X, Z] > -Z < Y, X > + < X, \nabla_Z Y >= \\
&= X < Y, Z > - < Y, [X, Z] > -Z < Y, X > + \\
&+ < X, [Z, Y] > + < X, \nabla_Y Z >= \\
&= X < Y, Z > - < Y, [X, Z] > -Z < Y, X > + < X, [Z, Y] > + \\
&+ Y < X, Z > - < \nabla_Y X, Z >= \\
&= X < Y, Z > - < Y, [X, Z] > -Z < Y, X > + < X, [Z, Y] > + \\
&+ Y < X, Z > - < [YX], Z > - < \nabla_X Y, Z >
\end{aligned}$$

Hence

$$\begin{aligned}
2 < \nabla_X Y, Z > &= X < Y, Z > -Z < Y, X > +Y < X, Z > - \\
&- < [YX], Z > - < Y, [X, Z] > + < X, [Z, Y] > .
\end{aligned}$$

This is known as Koszul's formula for the covariant derivative. Note that the bracket of two vector fields is uniquely determined on S by the inclusion $S \subset \mathbb{R}^3$.

$\square$

This shows that the covariant derivative of two vector fields only depends on the scalar product (and the derivatives) defined on the tangent space to S. In other words the covariant derivative is an intrinsic quantity (in spite of the fact that our original definition involves a derivative in $\mathbb{R}^3$)

Let S be a surface and let $X : U \to S$ be a local parametrization. Sometimes we will use the notation (u_1, u_2) (instead of (u, v)) for the coordinates in U, as this allows the use of summations. Moreover we let

$$X_1 = X_{u_1}, \qquad X_2 = X_{u_2}$$

Definition 2.16 *Let S be a surface and let $X : U \to S$ be a local parametrization of S. We define the Christoffel symbols of S (relative to the parametrization X) by*

$$\nabla_{X_i} X_j = \sum_k \Gamma_{ij}^k X_k$$

for $i, j = 1, 2$.

The vectors X_1, X_2 span, locally, the tangent space to S. Since the covariant derivative of two tangent vector fields is a tangent vector field we have that $\nabla_{X_i} X_j$ must be a linear combination of X_1 and X_2. Then Γ_{ij}^k are just the coefficients of this linear combination. For example

$$\nabla_{X_1} X_2 = \Gamma_{12}^1 X_1 + \Gamma_{12}^2 X_2.$$

Proposition 2.7 *Let $X : U \to S$ be a local parametrization of a surface S and let Γ_{ij}^k be the corresponding Christoffel symbols. Then $\Gamma_{ij}^k = \Gamma_{ji}^k$ and*

$$\Gamma_{ij}^m = \frac{1}{2} \sum_l g^{lm} (X_i g_{jl} + X_j g_{il} - X_l g_{ij}). \tag{2.5}$$

Proof: 1) Since X_i and X_j are coordinate vector fields we have

$$\begin{aligned} 0 &= [X_i, X_j] = \nabla_{X_i} X_j - \nabla_{X_j} X_i = \\ &= \sum_k \Gamma_{ij}^k X_k - \sum_k \Gamma_{ji}^k X_k = \\ &= \sum_k (\Gamma_{ij}^k - \Gamma_{ji}^k) X_k \end{aligned}$$

Since X_1 and X_2 are linearly independent the coefficients of this linear combination must vanish.

2) At each point of $X(U)$, the vectors X_1 and X_2 span the tangent space to S. This is an Euclidean vector space (with the scalar product induced by the scalar product in $\mathbb{R}^3$, i.e. the second fundamental form). Given three indices $i, j, l \in \{1, 2\}$, we have

$$\begin{aligned} X_l g_{ij} &= X_l < X_i, X_j > = < \nabla_{X_l} X_i, X_j > + < X_i, \nabla_{X_l} X_j > = \\ &= < \sum_k \Gamma_{li}^k X_k, X_j > + < X_i, \sum_k \Gamma_{lj}^k X_k > = \\ &= \sum_k \Gamma_{li}^k g_{kj} + \sum_k \Gamma_{lj}^k g_{ik} \end{aligned}$$

We can permute the indices and get two other relations

$$X_i g_{jl} = \sum_k \Gamma_{ij}^k g_{kl} + \sum_k \Gamma_{il}^k g_{jk},$$

$$X_j g_{li} = \sum_k \Gamma_{jl}^k g_{ki} + \sum_k \Gamma_{ji}^k g_{lk}.$$

We sum these last two relations and we subtract the first one:

$$X_i g_{jl} + X_j g_{il} - X_l g_{ij} = 2\sum_k \Gamma_{ij}^k g_{kl}.$$

Now we multiply on both sides by g^{lm} and sum over l:

$$\sum_l g^{lm}(X_i g_{jl} + X_j g_{il} - X_l g_{ij}) = 2\sum_k \Gamma_{ij}^k \sum_l g_{kl} g^{lm}.$$

Note that

$$\sum_l g_{kl} g^{lm}$$

is the product of the k-th row of the matrix G and the m-th column of the matrix G^{-1}. Since $G \cdot G^{-1} = I$ we have

$$\sum_l g_{kl} g^{lm} = \delta_{km}$$

and the previous formula reduces to

$$\sum_l g^{lm}(X_i g_{jl} + X_j g_{il} - X_l g_{ij}) = 2\Gamma_{ij}^m.$$

□

Note that from 2) we get another (compare with Koszul's formula) proof of the fact that the covariant derivative (completely determined by the Christoffel symbols) only depends on the first fundamental form of S.

One of the crucial tools in the study of curves was the Frenet frame, a basis of $\mathbb{R}^3$ defined at every point of the curve. The way this frame changes from point to point can be used (via the Frenet equations) to study the geometry of the curve. We want something analogue for the case of surfaces. At every point of a local parametrization we have a well defined frame $X_1, X_2, N = X_1 \times X_2$ and we want to study the way this frame changes on S. This is done by studying the derivatives with respect to the two variables u_1 and u_2 that we use to define the parametrization (the analogue, for a curve $\alpha(t)$, was the derivative with respect to the parameter t). Here we introduce a different notation for the second fundamental form:

$$A_{11} = e = X_{uu} \cdot N, \quad A_{12} = A_{21} = f = X_{uv} \cdot N, A_{22} = g = X_{vv} \cdot N.$$

Proposition 2.8 *Let $X : U \to S$ be a local parametrization of a surface S. Then*

(i)

$$D_{X_i} X_j = \sum_k \Gamma_{ij}^k X_k + A_{ij} N, \tag{2.6}$$

(ii)

$$D_{X_i} N = -\sum_{jk} g^{jk} A_{ij} X_k. \tag{2.7}$$

Proof: 1) We have

$$\begin{aligned} D_{X_i} X_j &= \left(D_{X_i} X_j\right)^{tang} + \left(D_{X_i} X_j\right)^{norm} = \\ &= \nabla_{X_i} X_j + < D_{X_i} X_j, N > N = \\ &= \sum_k \Gamma_{ij}^k X_k + < \frac{\partial}{\partial u_i} \frac{\partial X}{\partial u_j}, N > N = \\ &= \sum_k \Gamma_{ij}^k X_k + < X_{ij}, N > N = \sum_k \Gamma_{ij}^k X_k + A_{ij} N. \end{aligned}$$

2) Since the norm of N is constant we have

$$0 = X_i < N, N >= 2 < D_{X_i} N, N >$$

hence $D_{X_i} N$ is orthogonal to the tangent vector, and must be a linear combination of the vectors X_1 and X_2. Since $< N, X_j >= 0$ is constant we also have

$$\begin{aligned} 0 &= X_i < N, X_j >=< D_{X_i} N, X_j > + < N, D_{X_i} X_j >= \\ &= < D_{X_i} N, X_j > + < N, \sum_k \Gamma_{ij}^k X_k + A_{ij} N >= \\ &= < D_{X_i} N, X_j > + < N, A_{ij} N >=< D_{X_i} N, X_j > + A_{ij}. \end{aligned}$$

and

$$< D_{X_i} N, X_j >= -A_{ij}.$$

It follows (see Example 2.26) that the coefficient of $D_{X_i}N$ as linear combination of X_1 and X_2 are given by

$$G^{-1} \cdot \begin{pmatrix} A_{i1} \\ A_{i2} \end{pmatrix}$$

or

$$D_{X_i}N = -\sum_{jk} g^{jk} A_{ij} X_k.$$

□

In the theory of curves we showed (see Theorem 1.3) that the curvature and the torsion can be prescribed and completely characterize a curve up to rigid motions. We would like to have a similar result for surfaces, where the role of the curvature and the torsion is taken by the two fundamental forms. In general the coefficients $g_i j$ of the first fundamental form and A_{ij} of the second fundamental form are subject to some extra condition:

Proposition 2.9 *Let S be a regular surface and let $X : U \to S$ be a local parametrization of S. Then the coefficients of the first and the second fundamental form are subject to the following compatibility conditions:*

$$X_k\left(\Gamma_{ij}^l\right) - X_j\left(\Gamma_{ik}^l\right) + \sum_p \left(\Gamma_{ij}^p \Gamma_{kp}^l - \Gamma_{ik}^p \Gamma_{jp}^l\right) = \sum_p g^{pl}\left(A_{ij}A_{kp} - A_{ik}A_{jp}\right)$$

for $i, j, k, l \in \{1, 2\}$, known as Gauss equations. And

$$X_k(A_{ij}) - X_j(A_{ik}) + \sum_l \left(\Gamma_{ij}^l A_{kl} - \Gamma_{ik}^l A_{jl}\right) = 0$$

for $i, j, k \in \{1, 2\}$, known as Codazzi-Mainardi equations.

Proof: We take the derivative, with respect to X_k of (2.6):

$$\begin{aligned}
D_{X_k}\left(D_{X_i}X_j\right) &= D_{X_k}\left(\sum_l \Gamma_{ij}^l X_l + A_{ij}N\right) = \\
&= \sum_l D_{X_k}\left(\Gamma_{ij}^l X_l\right) + X_k(A_{ij})N + A_{ij}D_{X_k}N = \\
&= \sum_l X_k\left(\Gamma_{ij}^l\right) X_l + \sum_l \Gamma_{ij}^l D_{X_k}X_l + X_k(A_{ij})N + A_{ij}D_{X_k}N = \\
&= \sum_l X_k\left(\Gamma_{ij}^l\right) X_l + \sum_l \Gamma_{ij}^l \left(\sum_p \Gamma_{kl}^p X_p + A_{kl}N\right) +
\end{aligned}$$

$$
\begin{aligned}
+ & \quad X_k(A_{ij})N - A_{ij}\sum_{pl} g^{pl}A_{kp}X_l = \\
= & \quad \sum_l \left(X_k\left(\Gamma_{ij}^l\right) + \sum_p \Gamma_{ij}^p\Gamma_{kp}^l - A_{ij}\sum_p g^{pl}A_{kp} \right) X_l + \\
+ & \quad \left(X_k(A_{ij}) + \sum_l \Gamma_{ij}^l A_{kl} \right) N
\end{aligned}
$$

(note that, in the last step we renamed some of the indices). We obtain similar formula switching the role of k and j:

$$
\begin{aligned}
D_{X_j}\left(D_{X_i}X_k\right) &= \sum_l \left(X_j\left(\Gamma_{ik}^l\right) + \sum_p \Gamma_{ik}^p\Gamma_{jp}^l - A_{ik}\sum_p g^{pl}A_{jp} \right) X_l + \\
&+ \left(X_j(A_{ik}) + \sum_l \Gamma_{ik}^l A_{jl} \right) N.
\end{aligned}
$$

But our fields are coordinate fields hence

$$
D_{X_k}\left(D_{X_i}X_j\right) = \frac{\partial}{\partial x_k}\frac{\partial}{\partial x_i}\frac{\partial X}{\partial x_j} = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i}\frac{\partial X}{\partial x_k} = D_{X_j}\left(D_{X_i}X_k\right)
$$

hence we can compare the normal and the tangential part of the corresponding expressions. For the tangential part:

$$
X_k\left(\Gamma_{ij}^l\right) + \sum_p \Gamma_{ij}^p\Gamma_{kp}^l - A_{ij}\sum_p g^{pl}A_{kp} = X_j\left(\Gamma_{ik}^l\right) + \sum_p \Gamma_{ik}^p\Gamma_{jp}^l - A_{ik}\sum_p g^{pl}A_{jp}
$$

or

$$
X_k\left(\Gamma_{ij}^l\right) - X_j\left(\Gamma_{ik}^l\right) + \sum_p \left(\Gamma_{ij}^p\Gamma_{kp}^l - \Gamma_{ik}^p\Gamma_{jp}^l\right) = \sum_p g^{pl}\left(A_{ij}A_{kp} - A_{ik}A_{jp}\right).
$$

For the normal part:

$$
X_k(A_{ij}) + \sum_l \Gamma_{ij}^l A_{kl} = X_j(A_{ik}) + \sum_l \Gamma_{ik}^l A_{jl}
$$

or

$$
X_k(A_{ij}) - X_j(A_{ik}) + \sum_l \left(\Gamma_{ij}^l A_{kl} - \Gamma_{ik}^l A_{jl}\right) = 0.
$$

□

Then we have the following result (we omit the proof here)

Theorem 2.7 *[Bonnet] Let g_{ij}, A_{ij} be smooth functions defined in an open set $V \subset \mathbb{R}^2$, for $i, j = 1, 2$, such that the matrix $G = (g_{ij})$ is positive definite. Assume that these function satisfy the Gauss and Codazzi-Mainardi equations. Then for every $q \in V$ there exists a neighborhood $U \subset V$ of q and a diffeomorphism $X : U \to X(U) \subset \mathbb{R}^3$ such that $X(U)$ is a regular surface that has g_{ij} and A_{ij} as coefficients of the first and the second fundamental forms. Moreover, if U is connected and $\tilde{X} : U \to \tilde{X}(U)$ is another surface with the same fundamental forms then $X(U)$ and $\tilde{X}(U)$ differ by an isometry of $\mathbb{R}^3$.*

Another fundamental result of the theory of surfaces states that the Gauss curvature, that, locally, is defined by the ratio of the determinants of the two fundamental forms, is an intrinsic quantity. In other words K is completely determined by the first fundamental form (and its derivatives):

Theorem 2.8 *[Gauss-Theorema Egregium] Let S be a regular surface. Then the Gauss curvature K of S is invariant by local isometries.*

Proof: Using Corollary 2.1, it is enough to show that the Gauss curvature only depends on the first fundamental form. Consider a local parametrization $X : U \to S$ of S and the corresponding Gauss equation

$$X_k\left(\Gamma_{ij}^l\right) - X_j\left(\Gamma_{ik}^l\right) + \sum_p \left(\Gamma_{ij}^p\Gamma_{kp}^l - \Gamma_{ik}^p\Gamma_{jp}^l\right) = \sum_p g^{pl}\left(A_{ij}A_{kp} - A_{ik}A_{jp}\right)$$

and choose $i = j = 1$, $k = l = 2$, then

$$\begin{aligned} X_2\left(\Gamma_{11}^2\right) - X_1\left(\Gamma_{12}^2\right) + & \sum_p \left(\Gamma_{11}^p\Gamma_{2p}^2 - \Gamma_{12}^p\Gamma_{1p}^2\right) = \sum_p g^{p2}\left(A_{11}A_{2p} - A_{12}A_{1p}\right) = \\ = & \ g^{12}\left(A_{11}A_{21} - A_{12}A_{11}\right) + g^{22}\left(A_{11}A_{22} - A_{12}A_{12}\right) = \\ = & \ g^{22}\left(A_{11}A_{22} - A_{12}^2\right) \end{aligned}$$

where we used the fact that the second fundamental form is symmetric, i.e. $A_{ij} = A_{ji}$. Hence

$$A_{11}A_{22} - A_{12}^2 = \frac{1}{g^{22}}\left(X_2\left(\Gamma_{11}^2\right) - X_1\left(\Gamma_{12}^2\right) + \sum_p \left(\Gamma_{11}^p\Gamma_{2p}^2 - \Gamma_{12}^p\Gamma_{1p}^2\right)\right).$$

The left hand side is the determinant of the second fundamental form and the right hand side only depends on the first fundamental form. Hence we can conclude using the formula (2.3) for the Gauss curvature. $\square$

Example 2.27 *Let S be the xy plane in $\mathbb{R}^3$ with the obvious parametrization: $X(u,v) = (u,v,0)$. Then*

$$X_1 = (1,0,0), \quad X_2 = (0,1,0)$$

Since X_1 and X_2 are constant vectors, every derivative will vanish and we have $D_{X_i}X_j = 0$ for every $i,j \in \{1,2\}$. Hence $\Gamma_{ij}^k = 0$ $i,j,k \in \{1,2\}$.

Example 2.28 *Let S be a rotational surface parametrized by*

$$X(u,v) = (f(v)\cos(u), f(v)\sin(u), g(v)).$$

Then

$$X_1 = X_u = (-f(v)\sin(u), f(v)\cos(u), 0), \qquad X_2 = X_v = (f'(v)\cos(u), f'(v)\sin(u$$

Hence $g_{11} = E = f(v)^2$, $g_{12} = F = 0$ and $g_{22} = G = f'(v)^2 + g'(v)^2$. Then

$$G = \begin{pmatrix} f(v)^2 & 0 \\ 0 & f'(v)^2 + g'(v)^2 \end{pmatrix}, \qquad G^{-1} = \begin{pmatrix} \frac{1}{f(v)^2} & 0 \\ 0 & \frac{1}{f'(v)^2+g'(v)^2} \end{pmatrix}$$

and

$$\begin{aligned}
\Gamma_{11}^1 &= \frac{1}{2}\sum_l g^{l1}\left(X_1 g_{1l} + X_1 g_{1l} - X_l g_{11}\right) = \\
&= \frac{1}{2}\left(X_1 g_{11} + X_1 g_{11} - X_1 g_{11}\right) = \frac{1}{2}X_u(f(v)^2) = 0. \\
\Gamma_{11}^2 &= \frac{1}{2}\sum_l g^{l2}\left(X_1 g_{1l} + X_1 g_{1l} - X_l g_{11}\right) = \\
&= \frac{1}{2}g^{22}\left(X_1 g_{12} + X_1 g_{12} - X_2 g_{11}\right) = -\frac{f(v)f'(v)}{f'(v)^2 + g'(v)^2}. \\
\Gamma_{12}^1 &= \frac{1}{2}\sum_l g^{l1}\left(X_1 g_{2l} + X_2 g_{1l} - X_l g_{12}\right) = \frac{f'(v)}{f(v)}. \\
\Gamma_{12}^2 &= \frac{1}{2}\sum_l g^{l2}\left(X_1 g_{2l} + X_2 g_{1l} - X_l g_{12}\right) = 0. \\
\Gamma_{22}^1 &= \frac{1}{2}\sum_l g^{l1}\left(X_2 g_{2l} + X_2 g_{2l} - X_l g_{22}\right) = 0. \\
\Gamma_{22}^2 &= \frac{1}{2}\sum_l g^{l2}\left(X_2 g_{2l} + X_2 g_{2l} - X_l g_{22}\right) = \\
&= \frac{1}{2}g^{22}X_2 g_{22} = \frac{f'(v)f''(v) + g'(v)g''(v)}{f'(v)^2 + g'(v)^2}.
\end{aligned}$$

Parallel transport and geodesics

Given a vector v in $\mathbb{R}^n$ we can define a vector field V by letting $V(p) = v$ for every $p \in \mathbb{R}^n$. We can define the same vector field by saying that $V(p)$ and $V(q)$ only differ by a translation (to be precise $V(q)$ is the image of $V(p)$ under the differential of the translation that maps p to q). If $\alpha : I \to \mathbb{R}^n$ is a curve in $\mathbb{R}^n$ then the restriction $V(t)$ of V to the points of the trace of α is a constant vector field hence satisfies the differential equation $D_{\alpha'(t)}V(t) = \frac{d}{dt}V(t) = 0$, viceversa, a vector field $V(t)$ along α that satisfies that equation has constant components, hence $V(t_1)$ and $V(t_2)$ are parallel in $\mathbb{R}^n$ for every choice of t_1 and t_2. On a surface, in general, we do not have a notion of translation that can help us to define a field that is always parallel to a given vector, we will then use a differential equation:

Definition 2.17 *Let S be a surface and let $\alpha : I \to S$ be a regular curve. A smooth vector field V along α is parallel along α if*

$$\nabla_{\alpha'} V = 0$$

for every $t \in I$.

Proposition 2.10 *Let S be a surface an let $\alpha : I \to S$ be a regular curve. If the vector fields V_1 and V_2 are parallel along α then $< V_1, V_2 >$ is constant along α.*

Proof: We show that $< V_1, V_2 >$ is constant by taking the derivative with respect to the parameter t of the geodesic, this is the same as

$$\alpha' < V_1, V_2 > = < \nabla_{\alpha'} V_1, V_2 > + < V_1, \nabla_{\alpha'} V_2 > = 0.$$

$\square$

As an immediate consequence we have that

(i) If V is parallel along α then the norm of V is constant along α,

(ii) If V_1, V_2 are parallel along α then the angle between V_1 and V_2 is constant along α.

Theorem 2.9 *Let S be a surface and let $p \in S$, $v \in T_pS$. If $\alpha : I \to S$ is a regular curve such that $\alpha(0) = p$ then there exists a unique parallel vector field $V(t)$ along α such that $V(0) = v$.*

Proof: Let $X : U \to S$ be a local parametrization and suppose first that the trace of α is contained in $X(U)$. If $\alpha(t) = X(u_1(t), u_2(t))$ we can write

$$\alpha'(t) = u_1'(t)X_1 + u_2'(t)X_2, \qquad V(t) = a_1(t)X_1 + a_2(t)X_2$$

for some smooth functions $a_1(t), a_2(t)$. Then

$$\begin{aligned}
\nabla_{\alpha'}V &= \nabla_{u_1'(t)X_1+u_2'(t)X_2}V = u_1'(t)\nabla_{X_1}V + u_2'(t)\nabla_{X_2}V = \\
&= \sum_i u_i'(t)\nabla_{X_i}V = \sum_i u_i'(t)\nabla_{X_i}(a_1X_1 + a_2X_2) = \sum_{ij} u_i'\nabla_{X_i}\left(a_jX_j\right) = \\
&= \sum_{ij} u_i'\left(X_i(a_j)X_j + a_j\nabla_{X_i}X_j\right) = \\
&= \sum_{ij} u_i'X_i(a_j)X_j + \sum_{ij} a_j \sum_k \Gamma_{ij}^k X_k = \\
&= \sum_k \sum_i u_i'X_i(a_k)X_k + \sum_k \left(\sum_{ij} a_j u_i'\Gamma_{ij}^k\right) X_k = \\
&= \sum_k \left(\frac{da_k}{dt} + \sum_{ij} a_j u_i'\Gamma_{ij}^k\right) X_k
\end{aligned}$$

where, in the last step, we used the chain rule to take the derivative of a_k with respect to t. Hence V is parallel if and only if

$$\frac{da_k}{dt} + \sum_{ij} a_j u_i'\Gamma_{ij}^k = 0 \tag{2.8}$$

holds for $k = 1, 2$. This is first order a linear ODE that can be solved uniquely given the initial condition $V(0) = (a_1(0)X_1 + a_2(0)X_2) = v$ and the solution is defined in the whole interval I where the curve α is defined. The definition of parallel vector fields along α does not involve a parametrization and this concept is independent on the parametrization we choose to verify that V is parallel along α. Hence if the trace of α is not contained in a parametrization and we want to define V at $\alpha(t)$, we can cover the image of $[0, t]$ under α (a compact set since the image of a compact set under a continuous map is compact) with a finite number of local parametrizations and the equation can be solved on each of them. Due to the uniqueness of the solution V is well defined at $\alpha(t)$. □

Example 2.29 *If S is the xy plane in $\mathbb{R}^3$ with the usual parametrization, then the Christoffel symbols are identically zero. Hence the equation* (2.8) *just says that the component of a parallel vector field along any curve must be constant and we find the usual notion of parallelism in $\mathbb{R}^2$.*

Definition 2.18 *Let S be a surface and let $p \in S$, $v \in T_pS$. Given a curve $\alpha : I \to S$ such that $\alpha(0) = p$, the unique parallel vector field V along α such that $V(0) = v$ is called parallel transport of v along α.*

Example 2.30 *Let S be the torus parametrized by*

$$X(u, v) = ((R + r\cos(v))\cos(u), (R + r\cos(v))\sin(u), r\sin(v))$$

it is possibile to compute the Christoffel symbols, the only non-zero ones are:

$$\Gamma^2_{11} = \frac{1}{r}\sin(v)(R + r\cos(v)),$$

$$\Gamma^1_{12} = \Gamma^1_{21} = -\frac{r\sin(v)}{R + r\cos(v)}.$$

We will consider three curves on S and determine some parallel vector fields. Let $u_0 \in (0, 2\pi)$ and let $\alpha_1(t)$ be a coordinate curve $u = u_0$:

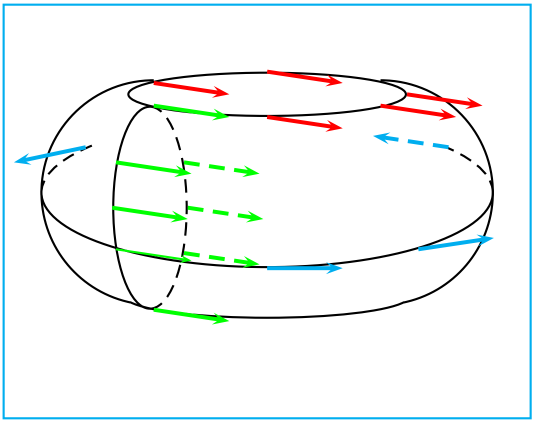

fig. 2.5

$$\alpha_1(t) = ((R + r\cos(t))\cos(u_0), (R + r\cos(t))\sin(u_0), r\sin(t))$$

the field X_u, restricted to α_1 is

$$X_u(t) = (-(R + r\cos(t))\sin(u_0), (R + r\cos(t))\cos(u_0), 0)$$

cannot be parallel since it does not have constant norm. We let

$$V_1(t) = \frac{1}{||X_u||}X_u = (-\sin(u_0), \cos(u_0), 0)$$

(green in fig. 2.5) then

$$\nabla_{\alpha_1'} V_1 = \left(\frac{d}{dt}V_1(t)\right)^{tang} = 0$$

i.e. V_1 is parallel along α_1. Similarly if we consider the upper circle:

$$\alpha(2)(t) = (R\cos(t), R\sin(t), r)$$

any tangent field that has constant components is parallel (for example the red one in fig. 2.5). In both cases the fields have constant components as vectors in $\mathbb{R}^3$ so the derivative w.r. to the parameter of the curve is always 0, but, in general, *a field with constant components does not remain tangent to the surface. If we consider the central circle ($v = 0$)*

$$\alpha(3)(t) = ((R+r)\cos(t), (R+r)\sin(t), 0)$$

and we let $V_3 = \alpha_3' = X_u$ (the blue one in fig. 2.5) then V_3 is parallel along α_3, we have

$$\alpha_3' = 1\,X_u + 0\,X_v, \qquad V_3 = 1\,X_u + 0\,X_v$$

hence we have $u_1' = 1, u_2' = 0$ and $a_1 = 1, a_2 = 0$. The system of ODE requires then $\Gamma_{11}^k = 0$ for $k = 1, 2$. Since $\Gamma_{11}^2 = 0$ for $v = 0$ we have that V_3 is parallel. This also shows that the tangent vector field to a curve $v = v_0$ is parallel if and only if $v = 0$ or $v = \pi$.

Remark 2.3

(i) *We used a local coordinate system to derive an equation for parallel vector fields V along a curve α, but the fact that V is parallel along α is independent on the chosen parametrization.*

(ii) *A vector field $V(t)$ restricted to a curve α on a surface S is parallel if $V'(t)$ does not have a component tangent to S. Then if $N(t)$ is the restriction of a normal field to S to the points of α, we have $V'(t) = \lambda(t)\,N(t)$ for some function $\lambda(t)$.*

(iii) *If $\phi : S_1 \to S_2$ is a local isometry then a vector field $V(t)$ is parallel along $\alpha : I \to S_1$ if and only if the vector field $d\phi(\alpha(t))(V)$ is parallel along $\beta(t) = \phi \circ \alpha(t)$. This follows from the fact that we have local parametrizations $X_1 : U \to S_1$ and $X_2 : U \to S_2$ such that the Christoffel symbols for S_1 and S_2 are the same at the corresponding point and, if we write $\alpha(t) = X_1(\gamma(t))$, $\beta(t) = X_2(\gamma(t))$ then the ODE for parallel vector fields along the two curves is the same.*

Proposition 2.11 *If two regular surfaces S_1 and S_2 are tangent along a curve $\alpha : I \to S_1 \cap S_2$ (i.e. $T_{\alpha(t)}S_1 = T_{\alpha(t)}S_2$) then a tangent vector field $V(t)$ is parallel along α on S_1 if and only if it is parallel along α on S_2.*

Proof: This is an immediate consequence of the previous remark. In fact if S_1 and S_2 are tangent along $\alpha(t)$ and $V(t)$ is parallel on S_1 then $V'(t) = \lambda(t)\,N_1(t)$, where $N_1(t)$ is normal to S_1. But then N_1 is also normal to S_2 and V is parallel on S_2. □

Example 2.31 *Let us consider the fields defined in the Example 2.30. The curve α_2 is tangent to both S and a plane, hence a vector field along α_2 is parallel if* *and only if it has constant components as vector in $\mathbb{R}^3$. The curve α_3 lies also on a cylinder tangent to the torus. The cylinder is locally isometric to the plane, and the pre-image of α_3 is a segment in $\mathbb{R}^2$. Since the tangent vector to a line is parallel in $\mathbb{R}^2$ we can conclude that V_3 is parallel on S.*

Example 2.32 *From the previous example it follows immediately that the tangent vector to a great circle on a sphere S^2 is parallel along the great circle: given a great circle we have a cylinder tangent to the sphere along that circle.*

Example 2.33

Let S be a regular surface and let $\alpha : I \to S$ be a regular curve. Once we have a nonzero parallel vector field V along α, it is easy to construct all other parallel vector fields:

(i) *Since $V(t)$ is parallel then $||V(t)||$ is constant and nonzero. Let $W(t)$ be a smooth unit vector field tangent to S and orthogonal to $V(t)$.*

(ii) *Let $t_0 \in I$ and let $Z(t)$ be the unique parallel vector field along α such that $Z(t_0) = W(t_0)$. Then, since the parallel transport preserves the length and the angles of parallel vector fields, we have that $||Z(t)|| = ||Z(t_0)|| = ||W(t_0)|| = 1$ and $Z(t)$ is orthogonal to $V(t)$. It follows that $Z(t) = W(t)$ and $W(t)$ is also parallel along α.*

(iii) *Let $v = aV(t_0) + bW(t_0) \in TS_{\alpha(t_0)}$ be a tangent vector to S. Define the field $U(t) = aV(t) + bW(t)$. Then*

$$\nabla_{\alpha'(t)} U(t) = \nabla_{\alpha'(t)}(aV(t) + bW(t)) = a\nabla_{\alpha'(t)} V(t) + b\nabla_{\alpha'(t)} W(t) = 0$$

hence $U(t)$ is the unique parallel vector field along α such that $U(t_0) = v$.

Definition 2.19 *Let S be a regular surface and let $\alpha : I \to S$ be a regular curve. Then α is a geodesic if $\alpha'(t)$ is a parallel vector field along α.*

From the definition it follows immediately that, if α is a geodesic, then $||\alpha'||$ is constant i.e. α is parametrized with constant speed.

Theorem 2.10 *Let S be a surface and let $p \in S$, $v \in T_pS$. Then there exists a unique geodesic $\alpha : (-\epsilon, \epsilon) \to S$ such that $\epsilon > 0$, $\alpha(0) = p$ and $\alpha'(0) = v$.*

Proof: This is an immediate consequence of the existence result for parallel vector fields. In a local coordinate system we have, with the notations of (2.8), $v_1 = u_1'$ and $v_2 = u_2'$ hence α is a geodesic if and only if:

$$u_k'' + \sum_{ij} u_i' u_j' \Gamma_{ij}^k = 0 \tag{2.9}$$

holds for $k = 1, 2$. This is first order a nonlinear ODE that can be solved uniquely given the initial condition $v(0)$, the solution is only guaranteed to exist in an interval centered in 0. □

Example 2.34 *Let $X(u_1, u_2) = (u_1, u_2, 0)$, $u_1, u_2 \in \mathbb{R}$, be a parametrization of the xy plane S in $\mathbb{R}^3$. The Christoffel symbols vanish and $\alpha(t) = (u_1(t), u_2(t), 0)$ is a geodesic if and only if $u_1''(t) = u_2''(t) = 0$. If $p = (a, b, 0)$ and $v = (v_1, v_2, 0) \in T_pS$, the geodesic through p, tangent to v, is the straight line*

$$\alpha(t) = (a, b, 0) + t\,(v_1, v_2, 0).$$

In general it is very difficult to solve the equation (2.9) explicitly, we will describe a few examples where the geodesics can be found using geometric arguments:

Example 2.35 *We can describe the geodesics of the sphere S^2 without using the equation* (2.9). *Let $\alpha(t)$ be a geodesic on S^2, parametrized by arclength. From $||\alpha'(t)|| = 1$ it follows that $< \alpha'(t), \alpha''(t) >= 0$. Taking the derivative of $1 = ||\alpha(t)||^2 =< \alpha(t), \alpha(t) >$ we obtain $< \alpha(t), \alpha'(t) >= 0$ and taking one more derivative*

$$0 =< \alpha(t), \alpha'(t) >' =< \alpha(t)', \alpha'(t) > + < \alpha(t), \alpha''(t) >= 1+ < \alpha(t), \alpha''(t) >$$

so that $< \alpha(t), \alpha''(t) >= -1$. If α is a geodesic then $\alpha'(t)$ is a parallel vector field along α i.e. the covariant derivative $\nabla_{\alpha'}\alpha'$ vanishes. This is equivalent to the fact that $\alpha''(t)$ does not have a component along the tangent space to S^2. Since $N(t) = \alpha(t)$ is a unit normal field at the points of the curve

$$\alpha''(t) =< \alpha''(t), N(t) > N(t) = -N(t) = -\alpha(t)$$

and α is a solution of the differential equation $\alpha''(t) + \alpha(t) = 0$. It is easy to check that the unique solution such that $\alpha(0) = p \in \S^2$ and $\alpha'(0) = v \in T_pS^2$ is given by

$$\alpha(t) = \cos(t)\,p + \sin(t)\,v$$

and $\alpha(t)$ is a great circle on S^2.

Example 2.36 *Let $S = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1\}$ be a cylinder in $\mathbb{R}^3$ and let $\alpha(t) = (x(t), y(t), z(t))$ be a geodesic parametrized by arclength. Then $x(t)^2 + y(t)^2 = 1$ and*

$$xx' + yy' = 0$$

taking one more derivative we obtain

$$x'^2 + y'^2 + xx'' + yy'' = 0$$

since α is parametrized by arclength: ${x'}^2 + {y'}^2 + {z'}^2 = 1$ and

$$xx'' + yy'' = 1 - {z'}^2. \qquad (2.10)$$

Since α is a geodesic, α'' is parallel to the normal vector $(x(t), y(t), 0)$ at every point of α hence

$$x''(t) = lambda(t)\,x(t), \quad y''(t) = lambda(t)\,y(t), \quad z''(t) = 0$$

for some smooth function λ. We will describe the geodesic such that $\alpha(0) = p = (1, 0, 0)$ and $\alpha'(0) = v = (0, a, b) \in T_pS$. From the third equation it follows that $z(t) = b\,t$ and $z'(t) = b$, hence, from (2.10), using the fact that $||v|| = 1$:

$$xx'' + yy'' = 1 - b^2 = -a^2$$

then, since α lies on the cylinder,

$$-a^2 = \lambda(t)x(t)^2 + \lambda(t)y(t)^2 = \lambda(t)(x(t)^2 + y(t)^2) = \lambda(t)$$

and $\lambda(t) = -a^2$ is constant. Then we can integrate the ODE's:

$$x'' + ax = 0, \qquad y'' + ay = 0$$

with our initial conditions, and we obtain $x(t) = \cos(at)$, $y(t) = \sin(at)$, so that

$$\alpha(t) = (\cos(at), \sin(at), bt).$$

If $b = 0$ the geodesic is a round circle, if $a = 0$ a vertical straight line, otherwise $\alpha(t)$ is an helix.

Given a point $p \in S$ and $v \in T_pS$ we have a geodesic $\alpha(t)$ tangent to v at p. If we rescale the vector v, then the corresponding geodesic is just a reparametrization of α:

Proposition 2.12 *Let S be a regular surface, $p \in S$ and, for $v \in T_pS$, denote by $\alpha_v(t) : (-\epsilon, \epsilon) \to S$ the geodesic such that $\alpha(0) = p$ and $\alpha'(0) = v$. Then, for $\lambda \in \mathbb{R} \setminus \{0\}$,*

$$\alpha_{\lambda v}(t) = \alpha_v(\lambda t)$$

is the geodesic through p, defined in $(-\frac{\epsilon}{\lambda}, \frac{\epsilon}{\lambda})$, tangent to λv.

Proof: Let $\beta(t) = \alpha_v(\lambda t)$. Then $\beta(0) = p$ and $\beta'(0) = \lambda$ $alpha'_v(0) = \lambda v$

$$\nabla_{\beta'}\beta' = \nabla_{\lambda\alpha'_v}(\lambda\alpha'_v) = \lambda^2\nabla_{\alpha'_v}\alpha'_v = 0$$

hence β is a geodesic. From the uniqueness of the solutions of the ODE (2.9) it follows that $\alpha_{\lambda v}(t) = \beta(t)$. $\square$

Corollary 2.3 *Let S be a regular surface, $p \in S$ then there is a neighborhood V of $0 \in T_pS$ such that for every $v \in V$, the geodesic $\alpha_v(t)$ is well defined for $t = 1$.*

Proof: Let $v \in T_pS$ be a unit vector, let $\epsilon(v)$ be the maximal interval such that $\alpha_v(t)$ is well defined in $[-\epsilon(v), \epsilon(v))$, where $p = \alpha_v(0)$. Then $\epsilon(v)$ is a continuous function (follows from the continuous dependence on initial data in the ODE (2.9)) defined in the unit circle, i.e. in a compact set. Let a be the minimum of $\epsilon(v)$. Then every geodesic spreading from p with unit speed is defined in $(-a, a)$. It follows that if $\delta < \frac{1}{a}$ and $||v|| < \delta$ the geodesic $\alpha_v(t)$ is well defined in $I = (-\frac{a}{\delta}, \frac{a}{\delta})$ and $1 \in I$. □

Definition 2.20 *Let S be a regular surface, $p \in S$ and let V be a neighborhood of $0 \in T_pS$ such that for every $v \in V$, the geodesic $\alpha_v(t)$ is well defined for $t = 1$. Then we define the exponential map at p: $exp_p : V \to S$ by*

$$exp_p(v) = \alpha_v(1), \qquad exp_p(0) = p.$$

Theorem 2.11 *Let S be a regular surface, $p \in S$. Then the exponential map at p: $exp_p : V \to S$ is a local diffeomorphism from a neighborhood of the origin in T_pS in S.*

Proof: The regularity of exp_p follows from results on the dependence of solutions of ODE's from initial data. We want to use the inverse function theorem to show that exp_p is a local diffeomorphism. For this purpose we have to evaluate the differential of exp_p at $0 \in V$. We identify the tangent space at 0 to T_pS with T_pS itself. Let v be a tangent vector at $0 \in V$ then $\beta(t) = tv$ is a curve in V that is tangent to v at $t = 0$. The image of $\beta(t)$ is the curve

$$exp_p(\beta(t)) = exp_p(tv) = \alpha_{tv}(1) = \alpha_v(t)$$

i.e. the geodesic starting from p with speed v (in the last equality we used Proposition 2.12). By definition

$$dexp_p(0)(v) = \frac{d}{dt}\alpha_v(t)_{t=0} = v$$

and the differential of exp_p at 0 is the identity map. This is clearly an invertibile linear map and the result follows. □

Since $exp_p : V \to S$ is a local diffeomorphism, defined in a neighborhood V of $0 \in T_pS$, we can use it to parametrize a neighborhood of p in S. We can parametrize V using the standard cartesian coordinate, then the corresponding coordinates (u, v) on for the points of S are called normal coordinates, while if we use polar coordinates in V then the coordinates in S are called polar geodesic coordinates.

If we use the polar geodesic coordinates (ρ, θ) in S, with $\rho \in (0, \epsilon)$, $\theta \in (0, 2\pi)$, then, by definition, the coordinate curves $\theta = \theta_0$ are the images of segments of straight lines through the origin in T_pS, i.e. geodesics in S. The images of the curves $\rho = \rho_0$ are called geodesic circles in S.

Proposition 2.13 *Let S be a regular surface and let $p \in S$. Let $p \in S$ and $V \subset T_pS$ such that $exp_p : V \to S$ is a diffeomorphism. Then, using the polar geodesic coordinates (ρ, θ) in S we have the following properties of the coefficients of the first fundamental form of S:*

$$E = 1, \qquad F = 0, \qquad \lim_{\rho \to 0} \sqrt{G} = 0, \qquad \lim_{\rho \to 0} (\sqrt{G})_\rho = 0.$$

Proof: The first statement is obvious since X_ρ is tangent to a unit speed geodesic. Since X_ρ and X_θ are coordinate vector fields we have $[X_\rho, X_\theta] = 0$ and

$$\begin{aligned} X_\rho(F) &= X_\rho < X_\rho, X_\theta >=< \nabla_{X_\rho} X_\rho, X_\theta > + < X_\rho, \nabla_{X_\rho} X_\theta >= \\ &= < X_\rho, \nabla_{X_\rho} X_\theta >=< X_\rho, -\nabla_{X_\theta} X_\rho >= \\ &= -\frac{1}{2} X_\theta < X_\rho, X_\rho >= -\frac{1}{2} X_\theta(1) = 0. \end{aligned}$$

Hence F does not depend of ρ we can compute it by taking the limit of $< X_\rho, X_\theta >$ as $\rho \to 0$. But it is easy to show that then $X_\theta \to 0$ (it is tangent to the geodesic circles) while X_ρ tends to a unit vector. Hence the limit of F is 0.

We omit the proof of the properties of G, it is possible to find it in the reference books. □

The fact that the coefficient F vanishes in polar geodesic coordinates is known as Gauss Lemma . This means that the radial geodesics from p meet the geodesic circles orthogonally, exactly like in the polar coordinates in an Euclidean space.

Proposition 2.14 *Let S be a regular surface and let $p \in S$. Let $p \in S$ and $V \subset T_pS$ such that $exp_p : V \to S$ is a diffeomorphism. Then, using the polar geodesic coordinates (ρ, θ) in S the Gauss curvature of S can be computed as*

$$K(\rho, \theta) = -\frac{(\sqrt{G})_{\rho\rho}}{\sqrt{G}}$$

Proof: We omit the proof of this result. The proof follows from the computation of the Christoffel symbols of S. □

We have then

Corollary 2.4 *[Minding] Two surfaces S_1, S_2 that have the same constant Gauss curvature are locally isometric.*

Proof: Let $p_1 \in S$ and $p_2 \in S$ and parametrize two neighborhoods of these points using polar geodesic coordinates, using the same range for the ρ and θ variables. From the previous results we have that the coefficients $E_1 = E_2 = 1$ and $F_1 = F_2 = 0$ of the first fundamental form are equal. The coefficients G_1 and G_2 satisfy the ODE

$$(\sqrt{G_i})_{\rho\rho} + K\sqrt{G_i} = 0.$$

Since K is constant we consider the following cases

1) $K = 0$. Then $(\sqrt{G_i})_{\rho\rho} = 0$ i.e. $(\sqrt{G_i})_\rho$ is a function $f_i(\theta)$ of the variable θ. We have

$$1 = \lim_{\rho\to 0}\sqrt{G_i})_\rho = \lim_{\rho\to 0} f_i(\theta)$$

and $f_i(\theta)$ is constant equal to 1. We can then integrate once more w.r to ρ:

$$\sqrt{G_i} = \rho + \tilde{f}_i(\theta)$$

for some function $\tilde{f}_i(\theta)$. But we also have

$$0 = \lim_{\rho\to 0}\sqrt{G_i}) = \lim_{\rho\to 0}\tilde{f}_i(\theta)$$

hence $\tilde{f}_i(\theta) = 0$ and $G_i = \rho^2$. The coefficients of the first fundamental forms are the same and the two surfaces are parametrized in the same neighborhood of the origin in an Euclidean space. Hence S_1 and S_2 are locally isometric.

2) $K > 0$. We can integrate w.r. to ρ and find

$$\sqrt{G_i} = A_i(\theta)\cos(\sqrt{K}\rho) + B_i(\theta)\sin(\sqrt{K}\rho)$$

for some functions A_i, B_i. Since $\sqrt{G} \to 0$ as $\rho \to 0$ we have $A_i(\theta) = 0$. Then

$$(\sqrt{G_i})_\rho = \sqrt{K}B_i(\theta)\cos(\sqrt{K}\rho)$$

and, since $(\sqrt{G})_\rho \to 0$ as $\rho \to 0$,

$$\sqrt{G_i} = \frac{1}{\sqrt{K}}\sin(\sqrt{K}\rho)$$

hence the coefficients of the first fundamental form are the same at the corresponding points and the two surfaces are locally isometric. We omit the proof for the case $K < 0$, which is similar. □

One of the important properties of geodesic is the following:

Proposition 2.15 *Let S be a regular surface and let $\beta(t)$ be a geodesic. Then β is locally length minimizing.*

Proof: Here we omit some of the details of the proof. Let p be a point on the trace of β and assume, for simplicity, that $p = \beta(0)$. Consider a normal geodesic neighborhood $\exp_p(V)$ of p and denote again by $\beta : [0, L] \to exp_p(V)$ a connected segment of β whose image lies entirely in $exp_p(V)$. Then $\beta(t) = \exp_p(tv)$ for some $v \in T_pS$. Let $q = \beta(L)$. And let $\alpha[0, L] \to S$ be another curve such that $\alpha(0) = p$, $\alpha(L) = q$ with $\alpha(t) \in exp_p(V)$ (we assume that the trace of α lies in the coordinate neighborhood to simplify the proof, in fact it is possible to show that this is not really necessary, up to a shrinking of V). We can parametrize $\alpha(t) : (0, L] \to exp_p(V) \subset S$ using polar geodesic coordinates $\alpha(t) = X(\rho(t), \theta(t))$, (where we assume that α does not pass through p twice, otherwise we consider just a component of α) then

$$< \alpha', \alpha' > = < \rho' X_\rho + \theta' X_\theta, \rho' X_\rho + \theta' X_\theta > = \rho'^2 + G\,\theta'^2 \geq \rho'^2$$

and the equality holds if and only if $\theta' = 0$. Then

$$L(\alpha) = \lim_{\epsilon \to 0} \int_\epsilon^L ||\alpha'(t)||\, dt \geq \lim_{\epsilon \to 0} \int_\epsilon^L |\rho'(t)|\, dt$$

on the right hand side we have the length of a curve such that $\theta = \theta_0$ is a constant, and, if $\rho' > 0$, this curve is the unique geodesic joining p and q. Hence β minimizes the distance between p and q. □

The converse is also true, any length minimizing regular curve α, parametrized by arclength, is a geodesic.

Example 2.37 *In a sphere a geodesic $exp_p(tv)$ remains length minimizing until it reaches the antipodal point $-p$. After that the minimizing geodesic is $exp_p(-tv)$ so the previous result cannot be extended to a global one.*

Let $X : V \to exp_p(V)$ be a normal geodesic neighborhood of p, and let α_1, α_2 be two different geodesics, parametrized by arclength. Then, for every fixed ρ_0, $\alpha_1(\rho_0)$ and $\alpha_2(\rho_0)$ belong to the same geodesic circle in S (note that a geodesic circle, in general, is not a geodesic) hence can be connected by a curve parametrized by $X(\rho_0, s)$, with $s \in [\theta_0, \theta_1]$. The length of this geodesic circle is then

$$L(\alpha_1(t), \alpha_2(t)) = \int_{\theta_1}^{\theta_2} ||X_\theta|| ds = \int_{\theta_1}^{\theta_2} \sqrt{G} ds.$$

We know something about the behavior of the coefficient $\sqrt{G}$, close to $\rho = 0$, we have that the initial slope at $\rho = 0$, is equal to 1, while the second derivative, with respect to ρ is equal to $-K\sqrt{G}$. Intuitively we see that the sign of the curvature controls the speed at which the two geodesic are spreading out of the point p. If $K < 0$ then the length of the geodesic circle connecting two points of the curves is increasing since $\sqrt{G}$ is a convex function. If $K = 0$ the length of the circle increases linearly. If $K > 0$ then $\sqrt{G}$ is a concave function and we have two possible behaviors: the length of the geodesic circle may continue to increase (this can be seen in a paraboloid, for the geodesics starting at the origin) or, at some point, it will start to decrease (e.g.in the sphere, where two great circles in p will meet at the antipodal point $-p$).

We have seen that the curvature of the manifold may force two geodesic that start at the same point to meet again at a different point (like in the case of the sphere). It follows that the exponential map cannot be a global diffeomorphism from the tangent space into the manifold, in fact it fails to be injective (points on two different lines through the origin have the same image) (this is also related to the fact that geodesic are only locally minimizing). Nevertheless, if the metric has some reasonable properties, the exponential map is defined in the whole tangent space:

Theorem 2.12 *[Hopf Rinow] Let S be a regular surface, then the following facts are equivalent*

(i) *S is a complete metric space.*
(ii) *any two points in S can be connected by a length minimizing geodesic.*
(iii) *the exponential map at a point $p \in S$ is defined in the whole tangent space T_pS.*
(iv) *every geodesic is defined for $t \in \mathbb{R}$.*

Recall that, using the first fundamental form, we can define the length of the curves in S, hence a distance in S. S is then a metric space and it is complete if and only if every Cauchy sequence is converging.

Corollary 2.5 *Every compact surface is complete*

Proof: Let $\alpha(t)$ be a geodesic parametrized by arclength such that $\alpha(0) = p$. Suppose that the geodesic is only defined in a maximal interval (on the right) of the form $[0, t_0)$. We want to prove that $t_0 = +\infty$. Given a sequence of points t_n in $[0, t_0)$ such that $t_n \to t_0$ we have that t_n is a Cauchy sequence in $\mathbb{R}$. Hence, given $\epsilon > 0$ we can find $n_0 \in \mathbb{N}$ such that

$$|t_n - t_m| < \epsilon, \qquad \forall m, n > n_0$$

since the geodesic is parametrized by arclength we have that the length of the arc α_n^m of α between $\alpha(t_n)$ and $\alpha(t_m)$ is $|t_n - t_m|$ hence, since α_n^m is one of the curves joining the two points

$$||\alpha(t_n) - \alpha(t_m)|| \leq L(\alpha_n^m) = |t_n - t_m| < \epsilon$$

if m, n are large enough. It follows that $\alpha(t_n)$ is a Cauchy sequence on S, the compactness of S implies that we can find a converging subsequence (that we still denote by $\alpha(t_n)$) such that $\alpha(t_n) \to p$. We let then $\alpha(t_0) = p$. This shows that we can extend α continuously to t_0, but it is possible to prove that this can be done in a smooth way. From the existence theorem for geodesic we have that there exists a geodesic in S such that $\gamma(t_0) = p$ and $\gamma'(t_0) = \alpha'(t_0)$ and γ is defined in $(-\epsilon, \epsilon)$. From the uniqueness part of that theorem we have that $\gamma = \alpha$. Hence α is defined in $[0, t_0 + \epsilon)$. Since t_0 was assumed to be maximal we have a contradiction unless $t_0 = +\infty$. The result now follows from the Hopf-Rinow theorem. $\square$

Let α be a curve (parametrized by arclength) on an oriented surface S. We have that α' is a unit vector tangent to S and $N(\alpha(t))$ is a unit normal vector. It follows that the vector field defined at the points of α by

$$U(t) = \alpha'(t) \times N(\alpha(t))$$

is a unit vector field tangent to S and orthogonal to α'. In other words

$$\{\alpha', N(\alpha(t)), U(t)\}$$

is an orthonormal frame at the points of $\alpha(t)$. It follows that we can write

$$\alpha''(t) = a(t)\alpha'(t) + b(t)N(\alpha(t)) + c(t)U(t)$$

but since $\alpha''(t) \perp \alpha'(t)$ (since the length of α' is constant) we have $a(t) = 0$. Moreover $\alpha''(t) \cdot N(\alpha(t))$ is by definition the normal curvature of α. Hence, using the symmetries of the triple product, we have

$$\alpha''(t) = k_n N + \alpha'' \cdot (\alpha' \times N) U = k_n N + N \cdot (\alpha'' \times \alpha') U = k_n N + |N||\alpha'' \times \alpha'| \cos(\theta) U$$

where θ is the angle between N and $\alpha'' \times \alpha'$. We define the geodesic curvature of α to be

$$k_g = |\alpha'' \times \alpha'| \cos(\theta)$$

so that

$$\alpha''(t) = k_n N + k_g U$$

(note that the sign of k_g depends on the choice of the orientation of S). Then the curvature of α is

$$k(t) = k_n(t)^2 + k_g(t)^2.$$

We have the following properties of the geodesic curvature:

(i) If α is a plane curve then $k_n = 0$, hence, up to sign, the curvature of α coincides with the geodesic curvature.

(ii) A curve, parametrized by arclength, on S is a geodesic if and only if $\alpha''(t)$ is a normal vector. This happens if and only if $k_g = 0$ since the geodesic curvature measures the length of the tangent part of α''. Hence the geodesic curvature measures the failure of α to be a geodesic.

Integration on surfaces

We give, without proofs, some results about the definition of integral of a function defined on a surface, i.e. given a function $f : S \to \mathbb{R}$, we want to define the integral of f over S. We try to find an analogue of the integral of a real valued function defined in an open subset U of $\mathbb{R}^n$. If $f : U \to \mathbb{R}$ is bounded we have the integral

$$\int_U f\, dx_1 \dots dx_n$$

and, if $\phi : U' \to U \subset \mathbb{R}^n$ is a diffeomorphism we have the change of variable formula:

$$\int_U f\, dx_1 \dots dx_n = \int U' f \circ \phi |\det(d\phi)|\, dx'_1 \dots dx'_n$$

the absolute value of the determinant of the Jacobian of ϕ has a geometric meaning that, in the case of two variables, corresponds to the area of the image of a square with unit sides under ϕ. In the case of a surface we have that a local parametrization plays the role of the diffeomorphism ϕ, mapping a two dimensional open subset of $\mathbb{R}^2$ into an open subset of a surface. In general it is not possible to cover the entire surface with just one coordinate neighborhood so we will have to find a way of making this local construction global.

Definition 2.21 *Let $X : U \to S$ be a local parametrization and let $f : S \to \mathbb{R}$ be a function such that the support of f, i.e. the closure of the set $\{p \in S : f(p) \neq 0\}$ is contained in $X(U)$. Then f is integrable if the following integral exists*

$$\int_U f \circ (X)(u,v)||X_u \times X_v||\, du\, d_v.$$

Here $X_u \times X_v$ is the area of the parallelogram spanned by X_u and X_v at each point of $X(U)$. This parallelogram is the image of a square spanned by $(1,,0)$ and $(0,1)$ in the tangent space at a point of U under the differential of X. Using the change of variable formula in $\mathbb{R}^2$ it is easy to show that this definition of integral does not depend on the particular choice of the parametrization. If the function f does not have support in $X(U)$ we have to use a partition of the unity:

Proposition 2.16 *Let S be a regular surface and let V_i be open subsets of S such that*

(i) $S = \bigcup V_i$
(ii) *every point of p has a neighborhood that interesects finitely many V_i*

Then there are functions $f_i : S \to \mathbb{R}$ such that

(i) *the support of f_i is contained in V_i*
(ii) $f_i(p) \geq 0$ *for every $p \in S$.*
(iii) *for every $p \in S$, $\sum f_i(p) = 1$*

we say that f_i form a partition of the unity subordinated to the covering V_i of S. It is possibile to prove that a partition of the unity always exists on S. Given a covering of S by coordinate neighborhoods $V_i = X_i(U_i)$ and function $f : S \to \mathbb{R}$ we have that $f \cdot f_i$ has support in V_i hence we can use the previous local definition of integral and set

$$\int_S f = \sum \int_{V_i} f\, f_i.$$

This formula remains valid if the covering of S leaves out a 'zero measure set'. By letting $f = 1$ the previous formula defines the area of a surface .

Example 2.38 *Let $X : (0, 2\pi) \times (-\frac{\pi}{2}, \frac{\pi}{2}) \to S^2$ be the local parametrization*

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

this parametrization covers the entire sphere with the exception of a half circle, this set does not contribute to the area of the sphere and we have

$$X_u \times X_v = (r^2\cos(v)^2\cos(u), r^2\cos(v)^2\sin(u), r^2\sin(v)\cos(v))$$

it follows that $||X_u \times X_v|| = r^2 \cos(v)^2$ *and we can compute the area of the sphere:*

$$\int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^2 \cos(v)\, du\, dv = \int_0^{2\pi} 2r^2\, du = 4\pi r^2.$$

The Gauss Bonnet theorem

The Gauss Bonnet theorem shows that there is a relation between the Gauss curvature of a surface and the topology of the surface. In the case of plane curves we see that it is possible, in some case, to associate a number, determined by the curvature, to a given plane curve, and this number is linked to the structure of the trace of the curve. Let $\alpha : [0, L] \to \mathbb{R}^2$ be a simple closed plane curve, i.e. $\alpha(0) = \alpha(L)$ and $\alpha(t_1) \neq \alpha(t_2)$ for $\{t_1, t_2\} \neq \{0, L\}$. It is possible to prove (Jordan's theorem) that α divides the plane in two regions, only one of them is bounded and we say that α is positively oriented if the normal vector points in the direction of the bounded region. Then we have

Theorem 2.13 *[Turning tangents] Let* $\alpha : [0, L] \to \mathbb{R}^2$ *be a simple, closed and positively oriented curve in* $\mathbb{R}^2$*. Then*

$$\int_0^L k\, dt = 2\pi.$$

Proof: Assume that α is parametrized by arclength. Then, if $\theta(t)$ is the angle between $\alpha'(t)$ and the positive direction of the x-axis, we have

$$\alpha'(t) = (\cos(\theta(t)), \sin(\theta(t)))$$

and

$$\alpha''(t) = \theta'(t)\,(-\sin(\theta(t)), \cos(\theta(t)))$$

our choice of the orientation implies that $\theta' > 0$ and we have:

$$k(t) = ||\alpha''(t)|| = \theta'(t)$$

hence

$$\int_0^L k\, dt = \int_0^L \theta'(t)\, dt = \theta(L) - \theta(0)$$

but, since the curve is simple and closed, the difference between the two angles must be 2π. $\square$

Given a simple closed curve it is possible to find a curve with the same trace that winds around several times (see (1.5)) in this case the integral of the curvature will be an integer multiple of 2π, i.e. by computing the integral of the curvature we may know how many times the curve winds around, this is a topological information about α (see the definition of fundamental group of a circle).

There is a non-smooth version of the previous result. If a curve $\alpha : I \to \mathbb{R}^2$ is only piecewise smooth and we have breakpoints for $t = t_1, \dots, t_n$, then we may define the exterior angle at t_i as the angle θ_i between the two vectors

$$\alpha'(t_i)_- = \lim_{t \to t_i -} \alpha'(t), \qquad \alpha'(t_i)_+ = \lim_{t \to t_i +} \alpha'(t)$$

if the curve is simple and closed then the integral of the curvature is less than 2π, but this can be compensated by adding the sum of the exterior angles:

$$\int_0^L k\, dt + \sum_i \theta_i = 2\pi.$$

Example 2.39 *Let α be the boundary of a triangle in $\mathbb{R}^2$. Then the sides of the triangle are straight line and the curvature is zero. The previous formula reduces to*

$$\sum_i \theta_i = 2\pi.$$

Since the exterior angles θ_i at the three vertices are related to the interior angles ϕ_i by $\theta_i = \pi - \phi_i$, it follows that $\sum \phi_i = \pi$, if we replace the triangle by a polygon with n sides, the same formula shows that the sum of the interior angles is $(n-2)\pi$. This facts can be proved by using elementary geometry but, using this formula, the proof is immediate. Note that the sides of a polygon are geodesics in $\mathbb{R}^2$.

There is a topological classification of the compact oriented surfaces in $\mathbb{R}^3$. Each such surface is characterized (up to diffeomorphism) by an integer, called genus of the surface. Intuitively the genus counts the number of 'holes' in the surface. For example the sphere has genus 0, the torus has genus 1... The genus is then a topological invariant of a surface and we can use it to define another topological invariant, called the Euler characteristic of the surface S :

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

where g is the genus of the surface. For example, the Euler characteristic of the sphere is 2, the one of the torus is 0.... The (first version) of the Gauss Bonnet theorem relates the Gauss curvature of the surface to the Euler characteristic in a very simple form:

Theorem 2.14 *[Gauss-Bonnet I] Let S be a compact orientable surface in $\mathbb{R}^3$. Then*

$$\int_S K = 2\pi\chi(S).$$

It follows that if we deform a round sphere, this deformation can create points where the curvature is not constant, positive or negative. But if the curvature is decreased in some area then there must be some other area where the curvature has increased since the total integral of the curvature is constant equal to 4π. This also shows that it is not possible to deform a torus so that the curvature is positive at all points, in fact the integral of the curvature remains equal to 0. We also have an important consequence

Corollary 2.6 *Let S be a compact oriented regular surface such that $K > 0$. Then S is homeomorphic to a sphere.*

Proof: It follows from the Gauss Bonnet theorem that the Euler characteristic of S is positive. Since $\chi(S) = 2 - 2g$ this is only possible if $g = 0$. □

Now we can state a different version of the Gauss Bonnet theorem, that applies to surfaces with boundary. We will not give a rigorous definition of surfaces with boundary. While a regular surface is locally diffeomorphic to an open connected sets in $\mathbb{R}^2$, we allow some coordinate neighborhoods in a surface with boundary to be modeled by open sets in the half plane $\{(x, y), y \geq 0\}$. The image of the boundary points of this set (i.e. the points on the x axis) will define the boundary of the surface. In most applications, for us, a surface with boundary is just the result of removing (open) regions bounded by closed simple curves on a regular surface:

Definition 2.22 *A regular region R in a regular surface S is a compact set given by the union of closures of connected disjoint open subsets of S such that the boundary of R is the union of piecewise smooth simple curves.*

We will assume that the boundary curves are 'positively oriented', this is an analogue of the choice of the orientation for simple closed. Intuitively, this means that if one is walking on the curve in the 'positive' direction and with one's head pointing to the normal vector N to the surface, then the region R remains to the left. It can be shown that one of the two possible orientations of the curve makes it positively oriented.

Theorem 2.15 *[Gauss-Bonnet II] Let S be a compact orientable surface in $\mathbb{R}^3$ and let R be a regular region whose boundary is the union of positively oriented piecewise smooth curves α_i. Denote by θ_{ij} the exterior angles at the singular points t_j for the curve α_i. Then*

$$\int_R K + \sum \int_{\alpha_i} k_g + \sum \theta_{ij} = 2\pi\chi(R)$$

where k_g denotes the geodesic curvature of the curve α_i.

This is an extremely powerful tool, but first one has to know how to compute the Euler characteristic of a regular region!

Definition 2.23 *A triangle in a regular surface S is a region T homeomorphic to a disc such that the boundary of T is a piecewise smooth curve that has three singular points with nonzero exterior angles. A triangulation of a region R in S is a family $\{T_i\}$ of triangles in S such that R is the union of the triangles and the intersection of two distinct triangles T_i and T_j can only be empty, a side (one of the smooth arcs in the boundary of the triangle) or a vertex (one singular point of the boundary).*

It is possible to prove that every regular region R in a regular surface admits a triangulation. Given a triangulation it is possible to compute the number F of faces (i.e. the triangles), the number E of the sides, and the number V of the vertices. Then we define the Euler characteristic of R

$$\chi(R) = F - E + V.$$

This quantity does not depend on the particular choice of the triangulation of R, hence it is a topological invariant of R.

Example 2.40 *Let R be the unit disc in the plane centered at the origin. We can divide it in 4 triangles defined by the intersections with the 4 quadrants. Then we have $F = 4$, $E = 8$, $V = 5$ (one vertex in the origin and four vertices on the boundary). Then $\chi(R) = 1$.*

Example 2.41 *We can divide each hemishpere of the sphere S^2 in 4 triangles that have a common vertex at a pole and the other vertices on a great circle (similar to the previous example). It follows that $F = 8$, $E = 12$, $V = 6$. Hence $\chi(S^2) = 2$.*

Example 2.42 *Let R be a truncated cylinder. We can divide it in two parts by cutting with a plane parallel to the axis, each of them is homeomorphic to a square, and we can subdivide it in two triangles. Then we have $F = 4$, $E = 8$ and $V = 4$. Hence $\chi(R) = 0$.*

Corollary 2.7 *Let S be a surface with $K \leq 0$. Then two distinct geodesic α_1, α_2 cannot intersect twice bounding a region R homeomorphic to a disc.*

Proof: We apply the Gauss Bonnet theorem to the region R. Since α_1 and α_2 are geodesic we do not have a contribution of the geodesic curvature and

$$\int_R K + \theta_1 + \theta_2 = 2\pi$$

since the geodesic cannot be mutually tangent, we have that the exterior angles are smaller than π. The fact that the integral gives a negative contribution leads to a contradiction. □

Corollary 2.8 *Let S be a surface, homeomprhic to a cylinder, with $K < 0$. Then S has at most one simple closed geodesic.*

Proof: Here we only give a trace of the proof. First one has to show that the two geodesic do not intersect (using the previous corollary). Then we have that the two geodesic bound a region homeomorphic to a cylinder R, we apply the Gauss Bonnet theorem to the region R and

$$\int_R K = 0$$

a contradiction, since $K < 0$. □

Given a triangle on a surface, assuming that the sides of the triangle are geodesics, one has from the Gauss Bonnet theorem:

$$\int_T K + \theta_1 + \theta_2 + \theta_3 = 2\pi$$

since the triangle is homeomorphic to a disc. If $\phi_i = \pi - \theta_i$ are the interior angles this implies

$$\phi_1 + \phi_2 + \phi_3 = \pi + \int_T K$$

hence the sum of the interior angles of a geodesic triangle in a sphere is larger than π.

We will now sketch the proof of the Gauss Bonnet I theorem. It is possible to show that, in a neighborhood of a point in S, we can find a local parametrization $X : U \to S$ such that the first fundamental form is of the form

$$I = \begin{pmatrix} \lambda(u,v) & 0 \\ 0 & \lambda(u,v) \end{pmatrix}$$

i.e. $E = G$ and $F = 0$. This has some geometric consequence: given $(u, v) \in U$ such that $X(u, v) = p \in X(U)$, the field $\frac{d}{du}(u, v) = (1, 0)$ and $\frac{d}{dv}(u, v) = (0, 1)$ form an orthonormal basis of the tangent space at (u, v) to $\mathbb{R}^2$. The differential of X maps these two fields into $X_u(p)$ and $X_v(p)$ and the first fundamental form determines how the differential of X has deformed this orthonormal basis. In the case of our parametrization we have that this deformation is just a rescaling by $\lambda(u, v)$, i.e. a similarity, and such a map preserves the angles between vectors. A map with this property is called conformal .

We will assume, for simplicity, that each triangle T in a triangulation of S is contained in the image of a conformal parametrization. Let $\alpha(t) : [0, L] \to S$ be the boundary of T and let $\tilde{\alpha}(t)$ be its pre-image under X, i.e. $\tilde{\alpha}(t) = (u(t), v(t))$ is a curve in $U \subset \mathbb{R}^2$ such that $\alpha(t) = X(u(t), v(t))$. The fact that the parametrization is conformal implies that the exterior angles θ_i at the singular points of α and $\tilde{\alpha}$ are the same. Furthermore a computation shows that the geodesic curvature k_g of α is related to the curvature $\tilde{k}$ of $\tilde{\alpha}$ by the formula

$$k_g = \frac{1}{2\lambda} (\lambda_u v' - \lambda_v u') + \tilde{k}.$$

The formula for the Gauss curvature in a conformal coordinate neighborhood has a particularly simple expression:

$$K = -\frac{1}{2\lambda} \Delta \log(\lambda)$$

and we can compute

$$\int_{\partial T} k_g + \sum_i \theta_i = \int_0^L \frac{1}{2\lambda} (\lambda_u v' - \lambda_v u') + \int_0^L \tilde{k} + \sum_i \theta_i$$

By the turning tangents theorem (and the fact that the parametrization is conformal) we have for the last two summands

$$\int_0^L \tilde{k} + \sum_i \theta_i = 2\pi.$$

Moreover by Green's theorem (the first integral on the right hand side can be seen as the integral of a function on the pre-image $\tilde{T}$ of T in $\mathbb{R}^2$)

$$\begin{aligned} \int_0^L \frac{1}{2\lambda} (\lambda_u v' - \lambda_v u') &= \int_{\tilde{T}} \frac{\partial}{\partial u} \left(\frac{\lambda_u}{2\lambda} \right) + \frac{\partial}{\partial v} \left(\frac{\lambda_v}{2\lambda} \right) du dv = \\ &= \int_{\tilde{T}} \frac{1}{2} \Delta(\log(\lambda)) du dv = \int_{\tilde{T}} \frac{1}{2\lambda} \Delta(\log(\lambda)) \lambda du dv = \\ &= -\int_T K \end{aligned}$$

since the determinant of the first fundamental form that appears in the formula for the integral of a function defined on a surface is, in this case, exactly λ. Hence we proved

$$\int_T K + \int_{\delta T} k_g + \sum_i \theta_i = 2\pi$$

if we sum over all the triangles T_i of the triangulation of R (note that this number is equal to the number of faces F in the triangulation, we get

$$\int_R K + \sum_j \int_{\delta T_j} k_g + \sum_j \sum_i \theta_i = 2\pi F$$

Now note that, since S has no boundary, each side of the triangles in the sum of the contribution of the geodesic curvatures, is counted twice, once for a triangle T_i and another one for a triangle T_k that shares that side with T_i. But the assumption on the orientation of the triangles implies that the two orientations, in T_i and T_k, are opposite. It follows that the two contributions will cancel and we are left with

$$\int_R K + \sum_j \sum_i \theta_i = 2\pi F$$

We have three exterior angles for each triangle hence

$$\sum_j \sum_i \theta_i = \sum_j \sum_i (\pi - \phi_i) = 3F\pi - \sum_j \sum_i \phi_i$$

but at each vertex the sum of the interior angles of the triangles that have that vertex in common is 2π hence

$$\sum_j \sum_i \theta_i = 3F\pi - 2\pi V$$

and

$$\int_R K = 2\pi F - 3\pi F + 2\pi V$$

it is possible to prove, by induction, that $3F = 2E$, hence

$$\int_R K = 2\pi F - 2\pi 2 + 2\pi V = 2\pi\chi(S).$$

Changing the metric

In the last section we studied the intrinsic geometry of a surface S in $\mathbb{R}^3$. In the formulas we computed quantities that only depended on the first fundamental form, i.e. the inner product induced on T_pS by the inclusion of S in $\mathbb{R}^3$. If we replace the first fundamental form by a different symmetric positive definite bilinear form on T_pS, the formulas for the curvature will still make sense. By doing this change we only allow some freedom in the way we measure the length of vectors in T_pS and this can make sense in some practical situations. Think of a map, this can be seen as a subset of $\mathbb{R}^2$, a flat object. But motion on the map is not equivalent to motion in the real space represented by the map, since, in general, the map represents a 'curved' piece of world and we cannot estimate the distance between points from their representatives on a map. This suggests that the length of a segment in the map should be measured differently according to the curvature of the space that it represents.

Definition 2.24 *Let S be a regular surface in $\mathbb{R}^3$. A Riemannian metric on S is defined by the assignment of a symmetric positive definite bilinear form $g(p)$ in the tangent space T_pS at each point of S. We require that g depends smoothly on the point, i.e. if X, Y are smooth vector fields on S, then $g(p)(X(p), Y(p))$ is a smooth function $S \to \mathbb{R}$.*

If we choose the usual basis $\{X_u, X_v\}$ of the tangent space at a point $p \in S$ given by a local parametrization, then $g(p)$ is still represented by a 2×2 matrix. We will then use then use the same notation used in the case of the inner product induced by standard product in $\mathbb{R}^3$:

$$E(p) = g(p)(X_u, X_u), \qquad F = g(p)(X_u, X_v), \qquad G(p) = g(p)(X_v, X_v)$$

and the formulas for the quantities defined by the first fundamental form, like the Christoffel symbols and the Gauss curvature still make sense. In the particular case $F = 0$ one can prove that the formula for the Gauss curvature takes the form

$$K = -\frac{1}{2\sqrt{EG}}\left(\frac{\partial}{\partial v}\left(\frac{E_v}{\sqrt{EG}}\right) + \frac{\partial}{\partial u}\left(\frac{G_u}{\sqrt{EG}}\right)\right) \tag{2.11}$$

and we will use it as the definition of the curvature of a Riemannian metric in the following examples.

Example 2.43 *Let S be the Poincare half-plane, defined by*

$$S = \{(x, y) \in \mathbb{R}^2 : y > 0\}$$

this is an open subset of $\mathbb{R}^2$, *hence with the induced metric this is just a flat space. But we can define a new metric, at point* $p = (u, v)$ *we let*

$$g(p) = \frac{1}{v^2} < \cdot, \cdot >$$

where $< \cdot, \cdot >$ *is the standard scalar product in* $\mathbb{R}^2$. *We are just rescaling the standard first fundamental point with a coefficient that depends on the point. At the points of the form* $(u, 1)$ *this is just the standard inner product of* $\mathbb{R}^2$, *but when* v *is large, the norm of a vector that has unit length in the Euclidean space, becomes very small. If* v *is close to zero then this length becomes very large. This means that the length of the tangent vector to a line segment of the form* $(0, t)$ *is such that the distance between* $(0, t_1)$ *and* $(0, t_2)$ *is huge if* t_1 *and* t_2 *are 'very small'. At every point of* S *we have the standard basis* $X_u = (1, 0)$ *and* $X_v = (1, 0)$. *It follows that*

$$E = G = \frac{1}{v^2}, \qquad F = 0$$

and we can use the formula (2.11) *to compute the curvature of* S

$$K = -\frac{v^2}{2}\left(\frac{\partial}{\partial v}\left(-\frac{2}{v^2}\right)\right) = -1$$

Example 2.44 *The stereographic projection* ϕ *from* $S = S^2 \setminus \{N\}$ *to* $\mathbb{R}^2$ *is a diffeomorphism. We can use this map to induce a metric on* S *by pretending that* ϕ *is an isometry, i.e. for* $X, Y \in T_pS$ *we let*

$$g(p)(X, Y) =< d\phi(p)(X), d\phi(p)(Y) >$$

where $< \cdot, \cdot >$ *is the standard scalar product in* $\mathbb{R}^2$. *This induces a new metric on an open subset of the sphere and this subset becomes isometric to the plane* $\mathbb{R}^2$, *hence the Gauss curvature of* S *is* 0.

It is possible to define the covariant derivative and the concept of geodesic using the formulas developed for the standard metric. Geodesic turn out to be again locally minimizers of the distance between two points:

Example 2.45 *Let* S *be the Poincare half-plane and let* $\gamma : [a, b] \to S$ *be a curve connecting two points* p_1, p_2 *that lie on the same vertical line,* $p_1 = (u_0, v_1)$, $p_2 = (u_0, v_2)$. *Then if we represent* γ *by* $(u(t), v(t))$,

$$\begin{aligned} L(\gamma &= \int_a^b \sqrt{g(\gamma', \gamma')}\, dt = \int_a^b \frac{1}{v(t)}\sqrt{u'(t)^2 + v'(t)^2}\, dt \geq \\ &\geq \int_a^b \frac{1}{v(t)}\sqrt{v'(t)^2}\, dt = \int_a^b \frac{|v'(t)|}{v(t)}\, dt \geq \int_a^b \frac{v'(t)}{v(t)}\, dt = \end{aligned}$$

$$= \int_{v_1}^{v_2} \frac{1}{v}\, dv = L(\alpha)$$

where $\alpha(t)$ is the line segment $[v_1, v_2] \to S$, defined by $\alpha(t) = (u_0, t)$. It follows that vertical line segments are geodesics. We can identify S with the subset of the complex numbers $\{z \in \mathbb{C} : Im(z) > 0\}$. Then it is possible to prove that the maps of the form

$$z \to \frac{az+b}{cz+d}$$

where $a, b, c, d \in \mathbb{R}$ and $ad - bc = 1$, are isometries of S. Since the image of a geodesic under an isometry is again a geodesic (since isometries preserve distance they must preserve length minimizing curves) one can consider the images of the geodesic $\alpha(t) = it$ (the positive v axis) and obtain the geodesics

$$t \to \frac{bd + act^2 + i}{d^2 + c^2t^2}.$$

It turns out that the geodesics are straight lines parallel to the v axis or half circles centered at a point of the u axis. For example for $a = b = c = 1$ and $d = 2$ one gets the circle of radius $\frac{1}{4}$ centered at $(\frac{3}{4}, 0)$.

The formula (2.5) that defines the Christoffel symbols in terms of the metric is still valid, we have the matrices

$$G = \begin{pmatrix} \frac{1}{v^2} & 0 \\ 0 & \frac{1}{v^2} \end{pmatrix}, \qquad G^{-1} = \begin{pmatrix} v^2 & 0 \\ 0 & v^2 \end{pmatrix}$$

and one can easily check that the only nonzero ones are

$$\Gamma^2_{11} = \frac{1}{v}, \qquad \Gamma^1_{12} = \Gamma^2_{21} = \Gamma^2_{22} = -\frac{1}{v}$$

and it follows that the differential equations satisfied by the geodesics in S are

$$u''v = 2u'v', \qquad v''v = (v')^2 - (u')^2$$

and the curves $(0, e^t)$ are clearly solutions. Here we see that the parametrization of the vertical straight lines is not the standard one!

Symbols

We list some symbols that are used in this text together with the page number where they first appear.

	Description	Pg.
$\underline{t} \times \underline{n}$	Cross product	14
A_{ij}	Components of the second fundamental form	61
$\underline{b}$	Bi-normal vector	14
$C^\infty(S)$	Smooth real valued functions in S	53
$d(x, y)$	Euclidean distance	2
$df(p)$	Differential of f at p	27
$D_X Y$	Derivative of a vector field	55
g_{ij}	Components of the first fundamental form G	57
$exp_p(v)$	Exponential map at p	73
g^{kl}	Components of the inverse of G	57
$H(p)$	Mean curvature of a surface	45
$H_p(f)$	Hessian of a function	51
$I_p(v, w)$	First fundamental form	41
$II_p(v, w)$	Second fundamental form	45
$k(t)$	Curvature of a curve	15
$k_n(t)$	Normal curvature of a curve	15
$K(p)$	Gauss curvature of a surface	45
$\underline{n}$	Normal vector	14
$N(p)$	Gauss map	44
$\nabla_X Y$	Covariant derivative of a vector field	55
$O(n)$	Orthogonal matrices	4
S_p	Weingarten map	45
$T_p S$	Tangent space at p to S	35
$T_v(x)$	Translation	3
$\underline{t}$	Tangent vector	13
$\tau(t)$	Torsion of a curve	15
$\|\|x\|\|$	Norm of x	2
$x \cdot y$	Scalar product	1
X_{12}	Change of coordinates	33
X_u	Partial derivative	41
$\mathfrak{X}(S)$	Vector fields in S	54

Bibliography

Do Carmo M.P. *Differential geometry of curves and surfaces*, Prentice Hall (1976)
Oprea J. *Differential geometry and its applications*, Prentice Hall (1997)
Shifrin T. *Differential geometry*, available online (2016)

Index