Differential Geometry of Curves and Surfaces

Luis Fernández
Department of Mathematics
University of Bath

Chapter 0

Introduction and preliminaries

The name of this course is *Differential Geometry of Curves and Surfaces*. Let us analyse each word to see what it is about.

Geometry is the part of mathematics that studies the 'shape' of objects. The name *geometry* comes from the greek *geo*, earth, and *metria*, measure; in the dawn of mathematics, geometry was the science of measuring the earth, which has been very important all along history (how can you sell or buy land if you cannot measure it?). It is, hence, one of the oldest branches of mathematics.

What do we mean by 'shape'? The Oxford English Dictionary defines shape as 'External form or contour; that quality of a material object (or geometrical figure) which depends on constant relations of position and proportionate distance among all the points composing its outline or its external surface; a particular variety of this quality.'

So shape is a quality that depends on distances among all points of the object. This implicity says that when we talk about the shape of an object we do not take into account where it is located in space: a bottle's shape is the same whether it is in the rubbish bin or on a table, since the distances and positions among all points of the bottle do not change.

For our purposes, shape will mean those characteristics of an object that are independent of position. In other words, if I take a surface, and I rotate it and/or translate it, the surface I obtain has the same shape as the original one. These movements of in space (rotations and translations) are called *rigid motions*.

In this course we will deal with curves living in the plane and in three-dimensional space as well as with surfaces living in three-dimensional space. A curve in space is essentially the shape that a wire would take. A surface is the shape that soap film, for example, takes.

It only remains to explain the word 'differential'. In order to measure the length of curves that live, say, in a surface, we need to give a meaning to the concept of velocity. Velocity is exactly what we express with the derivative (or 'differential'), as you may know from early calculus courses. So it means that in this course we will study the shape of curves making heavy use, among other things, of derivatives and concepts that are studied in calculus courses.

The curves and surfaces we will study are expressed with formulas; for example, a circle of radius 2 in the plane can be expressed as the set of points $(x, y) \in \mathbb{R}^2$ such that $x^2 + y^2 = 4$. Alternatively it can be expressed as the image of the map $\alpha : \mathbb{R} \to \mathbb{R}^2$ given by

$$\alpha(t) = (2\cos t, 2\sin t).$$

A sphere of radius 1 can be expressed as the set of points $(x, y, z) \in \mathbb{R}^3$ such that $x^2 + y^2 + z^2 = 1$.

Alternatively, it can be seen as the image of the map. $\sigma : \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$\sigma(\theta, \phi) = (\cos\phi\cos\theta, \cos\phi\sin\theta, \sin\phi).$$

These examples show that there are several ways to describe geometrical objects. We will use both ways, although most times the second, i.e. expressing the object as the image of a map, will be more suitable for calculations. This way of expressing geometric objects is called a 'parametrization', which gets this name from the fact that it introduces parameters (t in the example of the circle and ϕ and θ in the example of the sphere).

Since our geometric objects will live in the plane ($\mathbb{R}^2$) or in the space ($\mathbb{R}^3$), let us first review some facts about the vector space $\mathbb{R}^n$.

0.0.1 Review of geometry and linear algebra

The set $\mathbb{R}^n$ is the set of all n-tuples of real numbers. With the usual componentwise addition and multiplication by scalars it has the structure of a vector space, so its elements are called vectors. We will often also call them points, depending on whether we want to emphasise a direction (vectors, denoted like $\vec{v}$) or a location (points, denoted like $\boldsymbol{p}$). It should not be forgotten that both $\vec{v}$ and $\boldsymbol{p}$ are elements of $\mathbb{R}^n$. [Strictly speaking, *points* are elements of $\mathbb{R}^n$ when we think of it as *affine* space, which is something like a vector space not necessarily having a 0, i.e. without a reference point. $\mathbb{R}^n$ is an affine space too.]

Length, norms and scalar (dot) product

- The *length* or *norm* or *module* of a vector $\vec{v} = (v_1, v_2, \ldots, v_n)$ is defined as

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

- The *distance* between two points $\boldsymbol{p}$ and $\boldsymbol{q}$ in $\mathbb{R}^n$ is the module of their difference, i.e.

$$d(\boldsymbol{p}, \boldsymbol{q}) = \|\boldsymbol{p} - \boldsymbol{q}\|.$$

- The *scalar* (or *dot*) product of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ is

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

  The dot product is often also denoted as $\langle \vec{u}, \vec{v} \rangle$. Note that $\|\vec{v}\|^2 = \vec{v} \cdot \vec{v}$.

- Note that if we think of vectors in $\mathbb{R}^n$ as n by 1 matrices (i.e. column vectors), then $\vec{u} \cdot \vec{v} = \vec{u}^{\,t}\, \vec{v}$, where the ^{t} denotes 'transpose' and the product on the right term is matrix multiplication.
  If A is a matrix and we denote its column vectors as $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ (so that A could be written as $(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)$) then we have

$$A^t A = (\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)^t (\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) = \begin{pmatrix} (\vec{a}_1 \cdot \vec{a}_1) & (\vec{a}_1 \cdot \vec{a}_2) & \cdots & (\vec{a}_1 \cdot \vec{a}_n) \\ (\vec{a}_2 \cdot \vec{a}_1) & (\vec{a}_2 \cdot \vec{a}_2) & \cdots & (\vec{a}_2 \cdot \vec{a}_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\vec{a}_n \cdot \vec{a}_1) & (\vec{a}_n \cdot \vec{a}_2) & \cdots & (\vec{a}_n \cdot \vec{a}_n) \end{pmatrix}$$

- The scalar product satisfies $|\vec{u} \cdot \vec{v}| = \|\vec{u}\|\|\vec{v}\| \cos\theta$, where θ is the *angle* between $\vec{u}$ and $\vec{v}$.

- The *standard* basis of $\mathbb{R}^n$ will be denoted by $\{\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}\}$. Recall that $\vec{e_i}$ is an n-tuple of all 0's with a 1 in the i^{th} position. In $\mathbb{R}^3$ the notation $\{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ is also common and we may use it at times.

- A basis $\{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ is called *orthonormal* if $\vec{v_i} \cdot \vec{v_j} = \delta_{ij}$. (Recall that δ_{ij} is *Kronecker's delta*, i.e. a function with two integer entries (i and j) that takes the value 0 if $i \neq j$ and 1 if $i = j$.)

- If $\{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ is an orthonormal basis and $\vec{u}$ is any vector, then we can express $\vec{u}$ in terms of the basis as
$$\vec{u} = (\vec{u} \cdot \vec{v_1})\, \vec{v_1} + (\vec{u} \cdot \vec{v_2})\, \vec{v_2} + \cdots + (\vec{u} \cdot \vec{v_n})\, \vec{v_n}.$$

Vector (cross) product

The *vector* or *cross* product of two vectors, $\vec{u} = (u_1, u_2, u_3)$ and $\vec{v} = (v_1, v_2, v_3)$ in $\mathbb{R}^3$ is defined as

$$\vec{u} \times \vec{v} = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).$$

The vector product of two vectors is only defined in $\mathbb{R}^3$. It has the following properties.

- $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$

- $\vec{u} \times \vec{v} = \vec{0} \iff \vec{u}$ and $\vec{v}$ are linearly dependent (i.e. parallel or one of them is $\vec{0}$).

- $\vec{u} \times \vec{v}$ is perpendicular to both $\vec{u}$ and $\vec{v}$.

- $\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \|\vec{v}\| \sin\theta$ (θ is the angle between $\vec{u}$ and $\vec{v}$) and this quantity also gives the area of the parallelogram spanned by the vectors $\vec{u}$ and $\vec{v}$.

- $(\vec{u} \times \vec{v}) \cdot \vec{w}$ is the determinant of the matrix whose columns are the components of $\vec{u}$, $\vec{v}$ and $\vec{w}$ in this order. The absolute value of this number gives the volume of the parallelepiped spanned by $\vec{u}$, $\vec{v}$ and $\vec{w}$.

- An orthonormal basis $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ of $\mathbb{R}^3$ is *positively* oriented if $\vec{v}_3 = \vec{v}_1 \times \vec{v}_2$, and *negatively* oriented if $\vec{v}_3 = -\vec{v}_1 \times \vec{v}_2$ (note that these are the only two possibilities since $\vec{v}_3$ and $-\vec{v}_3$ are the only two unit vectors othogonal to both $\vec{v}_1$ and $\vec{v}_2$). Note that the standard basis of $\mathbb{R}^3$ is positively oriented.

Linear maps and matrices

A linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ is one that satisfies $L(a\vec{u} + b\vec{v}) = aL(\vec{u}) + bL(\vec{v})$ for $\vec{u}, \vec{v} \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$. In terms of bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, linear maps are represented by matrices.

Definition 0.0.1 *A matrix A is orthogonal if it is a square matrix that satisfies $A\,A^t = A^t\,A = I$, where the superscript ^{t} stands for the transpose and I is the identity matrix.*

A linear map $L : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if its matrix in terms of the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ is orthogonal.

If A is an orthogonal matrix, $\det(A)^2 = \det(A)\, \det(A^t) = \det(A\,A^t) = \det(I) = 1$, so $\det(A)$ is either 1 or -1. The set of orthogonal n by n matrices is denoted $O(n, \mathbb{R})$; the set of orthogonal n by n matrices with determinant 1 is denoted $SO(n, \mathbb{R})$.

If A is an orthogonal matrix,

- $(A\vec{u}) \cdot (A\vec{v}) = (\vec{u})^t A^t A \vec{v} = (\vec{u})^t \vec{v} = \vec{u} \cdot \vec{v}$. This implies that orthogonal maps leave lengths invariant: $\|A\vec{v}\| = \|\vec{v}\|$.

- $(A\vec{u}) \times (A\vec{v}) = \det(A)\, A\,(\vec{u} \times \vec{v})$. This can be proved as follows: if $\vec{w} \in \mathbb{R}^3$ is any vector,

$$\begin{aligned}
\vec{w} \cdot (A\vec{u}) \times (A\vec{v}) &= [\,(AA^{-1}\vec{w})] \cdot [(A\vec{u}) \times (A\vec{v})] \\
&= \det[A\,(A^t\vec{w}), A\vec{u}, A\vec{v}] \\
&= \det(A)\ \det(A^t\vec{w}, \vec{u}, \vec{v}) \\
&= \det(A)\ ((A^t\vec{w}) \cdot (\vec{u} \times \vec{v})) \\
&= \det(A)\ ((A^t\vec{w})^t(\vec{u} \times \vec{v})) \\
&= \det(A)\ (\vec{w}^t\, A\,(\vec{u} \times \vec{v})) \\
&= \det(A)\ (\vec{w} \cdot (A(\vec{u} \times \vec{v}))
\end{aligned}$$

  Reordering we have

$$\vec{w} \cdot [((A\vec{u}) \times (A\vec{v})) - \det(A)\ (A(\vec{u} \times \vec{v}))] = 0$$

  for any vector $\vec{w}$. Therefore,

$$((A\vec{u}) \times (A\vec{v})) - \det(A)\ A(\vec{u} \times \vec{v}) = 0,$$

  as desired.

- If $\{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ and $\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ are orthonormal bases of $\mathbb{R}^3$ then there is a unique orthogonal matrix A with $A\vec{u_1} = \vec{v_1}$, $A\vec{u_2} = \vec{v_2}$, $A\vec{u_3} = \vec{v_3}$. [In fact this matrix is, writing all vectors as columns in terms of the standard basis, $(\vec{v_1}, \vec{v_2}, \vec{v_3})\ (\vec{u_1}, \vec{u_2}, \vec{u_3})^t$.] This matrix has determinant 1 if both bases are have the same orientation (either positive or negative) and determinant -1 if they have different orientations. (This is in general for $\mathbb{R}^n$ although we only stated it for $\mathbb{R}^3$).

Definition 0.0.2

- *A* translation *in $\mathbb{R}^n$ is a map $T : \mathbb{R}^n \to \mathbb{R}^n$ of the form $T(\boldsymbol{p}) = \boldsymbol{p} + \vec{v}$ for some vector $\vec{v}$ in $\mathbb{R}^n$. It is just moving all points of $\mathbb{R}^n$ by the vector $\vec{v}$.*
- *A* rigid motion *in $\mathbb{R}^n$ is the composition of a translation and an orthogonal transformation of determinant 1.*
- *An* isometry *of $\mathbb{R}^n$ is a map from $\mathbb{R}^n$ to $\mathbb{R}^n$ that preserves distance between points. This is, G is an isometry if and only if $\|G(\boldsymbol{p}) - G(\boldsymbol{q})\| = \|\boldsymbol{p} - \boldsymbol{q}\|$.*

It is an important theorem that we will not prove that the only isometries of $\mathbb{R}^n$ are compositions of translations and orthogonal transformations.

0.0.2 Calculus

The most general definition of differentiability is the following.

Definition 0.0.3 *A map $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is* differentiable *at a point $\boldsymbol{p} \in U$ if there exists a linear function $Df_{\boldsymbol{p}} : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$\lim_{\vec{h} \to \vec{0}} \frac{f(\boldsymbol{p} + \vec{h}) - f(\boldsymbol{p}) - Df_{\boldsymbol{p}}(\vec{h})}{\|\vec{h}\|} = 0$$

The linear map $Df_{\boldsymbol{p}}$ is called the derivative *of f at $\boldsymbol{p}$. The matrix associated to $Df_{\boldsymbol{p}}$ when we write it in terms of the standard basis is denoted $Jf_{\boldsymbol{p}}$ (the* Jacobian matrix*).*

Recall that the partial derivative of a function at the point $\boldsymbol{p}$ with respect to its k^{th} entry, denoted $D_k f_{\boldsymbol{p}}$, is the derivative of the function with respect to that entry when we consider all the other entries as constants. If the k^{th} entry is called x_k, this is traditionally denoted as

$$\frac{\partial f}{\partial x_k}(\boldsymbol{p}) \quad \text{or} \quad \left.\frac{\partial f}{\partial x_k}\right|_{\boldsymbol{p}}.$$

Although you may not be used to it, the notation $D_k f_{\boldsymbol{p}}$ is generally better than the traditional one because it does not assume any predefined name of the variables. We had to assume that the k^{th} variable was called x_k in order to use the traditional notation. This often leads to errors when using the chain rule.

When f is a function of a single variable, we normally use the notation $f'(p)$ for the derivative of f at p.

The entry at position (ij) of the matrix $Jf_{\boldsymbol{p}}$ is given by (using traditional notation)

$$(Jf_{\boldsymbol{p}})_{ij} = \left.\frac{\partial f_i}{\partial x_j}\right|_{\boldsymbol{p}}.$$

Example 0.0.1 Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be a differentiable map. Let us use (u, v) for the variables of f, and let us use subindexes to denote the components of f so that

$$f(u, v) = (f_1(u, v), f_2(u, v), f_3(u, v)).$$

Let (p_1, p_2) be a point in $\mathbb{R}^2$. Then, using the traditional notation,

$$Jf_{(p_1,p_2)} = \begin{pmatrix} \left.\frac{\partial f_1}{\partial u}\right|_{(p_1,p_2)} & \left.\frac{\partial f_1}{\partial v}\right|_{(p_1,p_2)} \\ \left.\frac{\partial f_2}{\partial u}\right|_{(p_1,p_2)} & \left.\frac{\partial f_2}{\partial v}\right|_{(p_1,p_2)} \\ \left.\frac{\partial f_3}{\partial u}\right|_{(p_1,p_2)} & \left.\frac{\partial f_3}{\partial v}\right|_{(p_1,p_2)} \end{pmatrix}.$$

For example, if $f(u, v) = (u^2, v^2, uv)$ then

$$Jf_{(u,v)} = \begin{pmatrix} 2u & 0 \\ 0 & 2v \\ v & u \end{pmatrix}, \quad \text{so for example} \quad Jf_{(1,2)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \\ 2 & 1 \end{pmatrix}.$$

If you have trouble remembering where the functions and where the variables lie in the Jacobian matrix, maybe it is useful to remember that each **V**ariable takes a **V**ertical.

Properties of the derivative

In this course we will always assume that functions are C^∞ i.e. they are differentiable as many times as we want. Assuming this we have

- Mixed partials commute: $D_i D_j f_{\boldsymbol{p}} = D_j D_i f_{\boldsymbol{p}}$. Or in the classical notation, $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$.
- Product rule: If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are vector-valued functions of a single variable then

$$\begin{aligned} (\boldsymbol{X} \cdot \boldsymbol{Y})' &= \boldsymbol{X}' \cdot \boldsymbol{Y} + \boldsymbol{X} \cdot \boldsymbol{Y}' \\ (\boldsymbol{X} \times \boldsymbol{Y})' &= \boldsymbol{X}' \times \boldsymbol{Y} + \boldsymbol{X} \times \boldsymbol{Y}' \end{aligned}$$

- Chain rule: In terms of linear maps, $D(f \circ g)_{\boldsymbol{p}} = Df_{g(\boldsymbol{p})} \circ Dg_{\boldsymbol{p}}$. In terms of matrices, $J(f \circ g)_{\boldsymbol{p}} = Jf_{g(\boldsymbol{p})}\, Jg_{\boldsymbol{p}}$ (product of matrices in the last term). In terms of partial derivatives, in classical notation, using y's to denote the entries of f and x's to denote the entries of g, we have
$$\frac{\partial(f_i \circ g)}{\partial x_j}\Big|_{\boldsymbol{p}} = \sum_k \frac{\partial f_i}{\partial y_k}\Big|_{g(\boldsymbol{p})} \frac{\partial g_k}{\partial x_j}\Big|_{\boldsymbol{p}}.$$

An interpretation of the linear map $Df_{\boldsymbol{p}}$ is as follows. If we are travelling along a curve $\alpha(t)$ with velocity vector $\vec{v}$ at time t_0 (so $\alpha'(t_0) = \vec{v}$) then, if we suddenly get 'teletransported' via the map f we will be moving with velocity vector equal to

$$(f \circ \alpha)'(t_0) = Df_{\alpha(t_0)}(\alpha'(t_0)) = Df_{\alpha(t_0)}(\vec{v}).$$

So $Df_{\boldsymbol{p}}$ tells us how velocities in $\boldsymbol{p}$ transform when we apply the function f.

For the purposes of this course we will need the following.

Definition 0.0.4 *A map $f : \mathbb{R}^n \to \mathbb{R}^m$ is called* regular at a point $\boldsymbol{p}$ *if $Df_{\boldsymbol{p}}$ has maximal rank or, in terms of matrices, the matrix $Jf_{\boldsymbol{p}}$ has maximal rank. Recall that this means that, if $m \geq n$, the n columns of $Jf_{\boldsymbol{p}}$ are linearly independent, and if $m < n$, the m rows of $Jf_{\boldsymbol{p}}$ are linearly independent.*

A map f is said to be a regular *map if it is regular at all the points of its domain.*

We will be primarily interested in maps $f : \mathbb{R}^2 \to \mathbb{R}^3$. In this case, f will be regular at a point $\boldsymbol{p}$ if and only if the vectors $D_1 f_{\boldsymbol{p}}$ and $D_2 f_{\boldsymbol{p}}$ in $\mathbb{R}^3$ are linearly independent. Since two vectors in $\mathbb{R}^3$ are linearly independent if and only if their cross product is not zero, this can also be expressed as

$$D_1 f_{\boldsymbol{p}} \times D_2 f_{\boldsymbol{p}} \neq \vec{0}$$

or, in the traditional notation, using u and v as the variables of f,

$$\frac{\partial f}{\partial u}\Big|_{\boldsymbol{p}} \times \frac{\partial f}{\partial v}\Big|_{\boldsymbol{p}} \neq \vec{0}.$$

0.0.3 Topology

Topology is the branch of mathematics concerned with continuous functions and everything that does not change under continuous deformations. We only need a few definitions.

Definition 0.0.5

- *An* open *ball centered at $\boldsymbol{p}$ with radius r in $\mathbb{R}^n$ is a set of the form*
$$B_r(\boldsymbol{p}) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{p}\| < r\},$$
i.e. it consists of all points at a distance less than r from $\boldsymbol{p}$.
- *A subset U of $\mathbb{R}^n$ is called* open *if about every point of U there is an open ball containing the point that is completely contained in U.*
- *Given a subset M of $\mathbb{R}^n$, a subset V of M is called open if it has the form $V = M \cap U$, where U is an open subset of $\mathbb{R}^n$.*

This definition is quite technical, as you can see, and will be studied in more detail in further courses. You just need to think of an open set as one in which, given a point, we can get to this point from nearby points in *all possible directions* without leaving the set. For example, the interval $(0, 1]$ (i.e. 1 is included), is not open: we cannot reach 1 from the right without leaving the set.

Now we define continuity.

Definition 0.0.6 *Let M and N be subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. A map $f : M \to N$ is*

- Continuous at $\boldsymbol{p}$ *if*
$$\lim_{\boldsymbol{x}\to\boldsymbol{p}} f(\boldsymbol{x}) = f(\boldsymbol{p}).$$
(Roughly this means that if we get close to $\mathbf{p}$ while staying in M, then the value of f will be close to $f(\boldsymbol{p})$ regardless of the way in which we get close to $\boldsymbol{p}$)
If the map f is continuous at every point, then it is just called continuous*.*

- *A* homeomorphism *if*
  - *It is continuous.*
  - *It is bijective (i.e. one-to-one and onto).*
  - *Its inverse is also continuous.*

Two sets are called homeomorphic *if there exists a homeomorphism between them.*

Roughly speaking, two sets are homeomorphic if we can deform one into the other without breaking or glueing anything. For example, a circle and an ellipse are homeomorphic.

NOTE: These definitions are rather technical and difficult. We will only need a general idea of these concepts for the rest of the course. They are here for the sake of completeness—without them we would not be able to give a rigorous definition of surfaces, for example. But we do not want to spend much time on them, as you will learn all about topology in more advanced courses.

Chapter 1

Curves

1.1 Regular curves and parametrised curves

Definition 1.1.1 *A* parametrised curve *is a smooth map $\alpha : (a, b) \to \mathbb{R}^n$, where $-\infty \leq a < b \leq \infty$ (i.e. a can be $-\infty$ and/or b can be ∞).*

A parametrised curve α is called regular *if in addition the derivative of α is never 0.*

Recall that 'smooth' means 'infinitely differentiable', i.e. we can take as many derivatives of α as we desire. We could have defined parametrised curves without this smoothness condition, but then we would have extremely messy objects about which not many general facts can be said. So in general **all the parametrised curves in this course will be smooth (i.e. infinitely differentiable) maps**.

The (regularity) condition $\alpha' \neq 0$ in the definition of parametrised regular curve is added to avoid cusps and corners.

Note that the above is not the way we think of curves in life. We never think of curves as maps, rather as a set of points in space. This is why we put the modifier 'parametrised' in the previous definition. Following this idea, a curve would be more likely defined as follows:

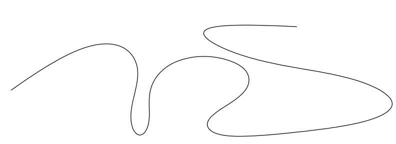
Definition 1.1.2 *A* regular curve *is a subset C of $\mathbb{R}^n$ such that, for every $p \in S$ there is an open set $U \subset \mathbb{R}^n$, with $p \in U$, and there is a smooth, one-to-one and onto map $\alpha : (a, b) \to C \cap U$ whose derivative is never 0.*

Let us unrevel the definition of a regular curve. It is a subset of space, which is intuitively the way we think of curves. The rest of the definition means that, locally (i.e. about any point of the curve), the curve is a parametrised regular curve as defined before.

Example 1.1.1

- A parametrised curve (not regular - look at $\alpha'(0)$): $(\frac{t^2}{t^2+1}, \frac{t^3}{t^2+1})$, $-\infty < t < \infty$.

- A regular curve:



Note that regular curves cannot have self-intersections: if they did, about the point of intersection the curve would look like an X, and there cannot be a one-to-one and onto smooth map from an interval to an X. Parametrised curves, instead, can have self intersections.

Example 1.1.2

- A regular parametrised curve with a self intersection: $(\sin t, \sin 2t)$, $-\infty < t < \infty$.

The difference between parametrized curves and regular curves is not very important, since every regular curve can be seen as the image of a parametrised curve, so our study will concentrate in parametrised curves.

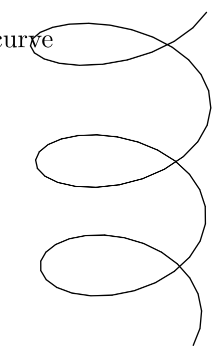
Note that if we have a parametrised curve α and we pre-compose it with a function ϕ from $\mathbb{R}$ to $\mathbb{R}$, we obtain a different parametrised curve $\tilde{\alpha} = \alpha \circ \phi$ with the same 'trace' as the original one. Then $\tilde{\alpha}$ will be called a *reparametrization* of α. For example, given the curve

$$\alpha(t) = (\cos t, \sin t, t), \quad \text{for} \quad -\infty < t < \infty$$

we can compose it with the function $\phi(s) = 3s$ to obtain the parametrised curve

$$\tilde{\alpha}(s) = (\cos 3s, \sin 3s, 3s).$$

The trace of both α and $\tilde{\alpha}$ is the same: a 'helix'.

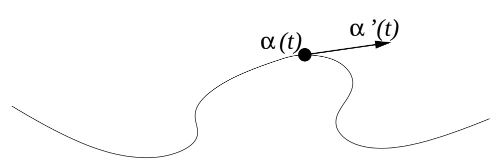


In our study of the 'shape' of curves we would like to think of curves with the same trace as equal. If the image of the maps is the same, why should we think of them as different? Hence all the properties that we will study will be **independent of reparametrization**.

Definition 1.1.3 *Let $\alpha : (a, b) \to \mathbb{R}^n$ be a parametrised curve, and let $\phi : (a', b') \to (a, b)$ be a differentiable bijective function whose derivative is never 0. The parametrised curve $\tilde{\alpha} : (a', b') \to \mathbb{R}^n$ defined by $\tilde{\alpha}(t) = \alpha(\phi(t))$ is called a* reparametrisation *of α.*

1.2 Velocity vector

Given a parametrised curve $\alpha : (a, b) \to \mathbb{R}^n$, its derivative $\alpha'(t)$ at some $t \in (a, b)$ will be a vector in $\mathbb{R}^n$ that gives the *velocity* of α at t. We can think of $\alpha'(t)$ as a vector anchored at the point $\alpha(t)$:



The tangent line to α at the point $\alpha(t_0)$ is the line

$$\ell(s) = \alpha(t_0) + s\alpha'(t_0).$$

While a reparametrisation of a curve leaves the trace of the curve invariant, what does it do to the velocity? Since the velocity is the derivative, we only have to take derivatives to $\tilde{\alpha}$ and apply the chain rule:

$$\tilde{\alpha}'(t) = \alpha'(\phi(t))\, \phi'(t).$$

Thus the velocity vector will be multiplied by $\phi'(t)$. In other words, if we are travelling along $\tilde{\alpha}(s)$, where we think of the parameter s as the time, we will do the same trajectory as if we travel along $\alpha(t)$ but we will be at different places in different moments, and our speed is likely to be different at each point of the curve. Note that we used different parameters t and s, related by $s = \phi(t)$, only to emphasise the fact that our time frame is different in each case.

Note that the reason why we defined 'regular' parametrised curves is that we would like to know what is a tangent direction to a curve. If α is 0 at some point then there is no way in general to define a tangent. Additionally it avoids cusps and corners:

(In a cusp or corner there is no well defined tangent, so either the derivative does not exist or it is zero at that point.)

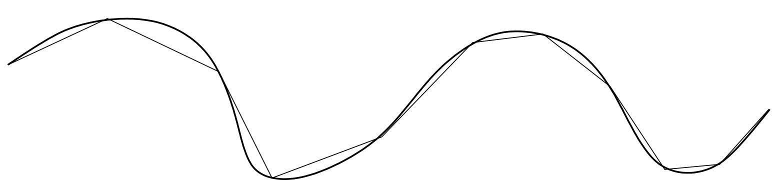
In the study of curves we want, as we said before, those properties independent of reparametrisation. Hence given a curve we can use a particular parametrisation to define the quantities that describe these curves. We can think of it as follows: if we want to compare how curved two curves in a road are we could just measure the centrifugal force as we drive through the curve. However, if we took each curve at different speeds our measurements would not be comparable, so we would not be able to conclude anything. Therefore, if we want to compare how curved curves are in a road we would have to take them all at the *same* speed.

When we study curves in $\mathbb{R}^n$ we would like to parametrise the curves so that the velocity vector has constant length at every point, and for convenience we choose this length to be 1. But is this always possible? We will soon see that it is, but we first need a definition.

1.3 Arc length

Given that we are studying geo*metry*, let us start measuring lengths of curves. We know intuitively how to measure the length of segments (and we have a definition for this: the length of the segment from $\boldsymbol{p}$ to $\boldsymbol{q}$ is $\|\boldsymbol{p} - \boldsymbol{q}\|$). How do we generalise this to not straight curves?

The first idea would be to approximate the curve with short segments and then measure the length of each segment and add the lengths up:



It is intuitively clear that, as we make the segments shorter and shorter we obtainy a better approximation to what we would like to call 'the length of the curve'. If we take the limit as the length of the segments goes to 0 (so we would be 'adding the lengths of infinitely many infinitesimally small segments') we get an integral. This motivates the following definition.

Definition 1.3.1 *The* length *of a parametrised curve $\alpha : (a, b) \to \mathbb{R}^n$ between the point t_0 and*

t_1 *is defined as*

$$\int_{t_0}^{t_1} \|\alpha'(t)\| dt.$$

This is the point in which it becomes clearer why the name of this course has the name 'differential' on it: we need to take derivatives in order to measure lengths. Actually the length of not differentiable curves (i.e. assuming that α is only continuous) can also be defined as the limit explained before the definition, provided that this limit exists. In this case the curve is called 'rectifiable'. There are, however, continuous curves for which this limit does not exist or equals infinity.

We can now show that parametrised regular curves can be reparametrised so that the velocity vector has length 1.

Proposition 1.3.1 *Every curve $\alpha : (a, b) \to \mathbb{R}^n$ can be reparametrised so that its velocity vector has length 1. This reparametrisation is called* parametrisation by arc length.

Proof : Let $t_0 \in (a, b)$ and consider the function

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

Since α is smooth (i.e. it can be differentiated as many times as we want), α' is also smooth. Since α' is never 0, $\|\alpha'\|$ is also smooth. Now, the integral of a smooth function is also smooth, so $s(t)$ is smooth.

The derivative of $s(t)$ is $s'(t) = \|\alpha'(t)\|$ by the fundamental theorem of calculus. [This is actually the crucial property of $s(t)$ and this is why the value of t_0 is irrelevant for the proof. In fact any function $s(t)$ that satisfies $s'(t) = \|\alpha'(t)\|$ will do the job.] Since the curve is regular, $\|\alpha'(t)\|$ is never 0, so it is either an always increasing or an always decreasing function, which implies that it is one-to-one and therefore it has a differentiable inverse.

Let $\phi(s)$ be the inverse of $s(t)$ (we are abusing notation here for convenience: note that s denotes both a variable and a function). Let $\tilde{\alpha}(s) = \alpha(\phi(s))$. Then, since $s(\phi(s)) = s$ we have, differentiating both sides with respect to s and using the chain rule, that $s'(\phi(s))\,\phi'(s) = 1$, and therefore,

$$\phi'(s) = \frac{1}{s'(\phi(s))} = \frac{1}{\|\alpha'(\phi(s))\|},$$

which always makes sense since α' is never 0.

Now we calculate $\tilde{\alpha}'$:

$$\tilde{\alpha}'(s) = \alpha'(\phi(s))\,\phi'(s) = \frac{\alpha'(\phi(s))}{\|\alpha'(\phi(s))\|}.$$

It is clear that the right hand side of the last expression has norm equal to 1. Therefore $\|\tilde{\alpha}'(s)\| = 1$ for all s.

□

Example 1.3.1 Let us find the parametrization by arc length of the curve $\alpha(t) = (e^t \cos t, e^t \sin t)$ (a spiral).

$$\alpha'(t) = (e^t \cos t - e^t \sin t, e^t \sin t + e^t \cos t).$$

$$\|\alpha'(t)\|^2 = (e^t \cos t - e^t \sin t)^2 + (e^t \sin t + e^t \cos t)^2 = 2e^{2t} \cos^2 t + 2e^{2t} \sin^2 t = 2e^{2t}.$$

Thus,

$$s(t) = \int \sqrt{2} e^t dt = \sqrt{2} e^t.$$

[Note that here we wrote an improper integral to define $s(t)$. As we said before we only need that $s(t)$ be an antiderivative of $\|\alpha'(t)\|$; for example, we could have also chosen $s(t) = \sqrt{2}e^t + C$, where C is some constant.] The inverse of $s(t)$ is $\phi(s) = \log(s/\sqrt{2})$, so a reparametrization by arc length of α is

$$\tilde{\alpha} = \left(\frac{s}{\sqrt{2}} \cos(\log(s/\sqrt{2})), \frac{s}{\sqrt{2}} \sin(\log(s/\sqrt{2})) \right).$$

Parametrisations by arc length are not unique, as we saw above. But they only depend on the 'origin of time' and the 'direction of time', as follows:

Lemma 1.3.1 *If α and β are two reparametrisations by arc length of the same parametrised curve γ, then either $\alpha(s) = \beta(s + c)$ or $\alpha(s) = \beta(-s + c)$, where the letter c stands for some constant. In the first case we say that the two parametrizations have the same orientation and in the second we say that they have opposite orientation (the curve is traced in opposite directions).*

Proof : Since α and β are reparametrisations of γ, we must have that α is also a reparametrisation of β, this is, $\alpha(s) = \beta(\phi(s))$ for some smooth bijective function ϕ. Taking derivatives in both sides and using the chain rule we obtain

$$\alpha'(s) = \beta'(\phi(s))\phi'(s),$$

and taking norms we find

$$1 = \|\alpha'(s)\| = \|\beta'(\phi(s))\| \, |\phi'(s)| = |\phi'(s)|.$$

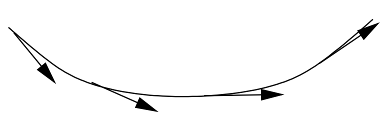
Hence either $\phi'(s) = 1$ for all s, which implies $\phi(s) = s + c$ for some constant c, and therefore $\alpha(s) = \beta(s + c)$ or $\phi'(s) = -1$ for all s, which implies $\phi(s) = -s + c$ for some constant c, and therefore $\alpha(s) = \beta(-s + c)$. Note that these are the only cases: $\phi'(s)$ cannot be 1 for some values of s and -1 for others since it is a continuous function.

□

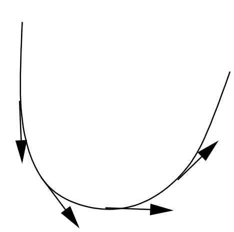
Great! So now we can start measuring things in curves without worrying about whether what we define is dependent on speed or not: we just take everything to have speed 1 (in these notes, speed will mean the length of the velocity vector, as it is usually understood).

1.4 Curvature

Our next goal is to make mathematical sense to the question 'How curved is a curve?' As we discussed before, when we are riding in a car we would say that, at same speed, the more curved is a curve, the more we feel the centrifugal force. The centrifugal ("running away from the centre") force is a reaction of the centripetal force, which is a measure of the rate of change of direcion of the velocity vector: the faster it changes, the stronger will be the centrifugal force we feel:



Velocity vector changes slowly — weak centrifugal force



Velocity vector changes rapidly — strong centrifugal force

Definition 1.4.1 *Let $\alpha : (a,b) \to \mathbb{R}^n$ be a parametrised curve, parametrised by arc length (i.e. $\|\alpha'(s)\| = 1$ for all s). The* curvature *of α at the point $\alpha(s)$ is the function*

$$\kappa(s) = \|\alpha''(s)\|.$$

It is traditionally denoted by the greek letter κ (kappa).

The curvature of a parametrised curve in general is defined as the curvature of any reparametrisation by arc length of that curve.

Of course this definition only makes sense if the number $\|\alpha''(s)\|$ is independent of the arc length parametrization of α. But it is easy to see that it is the case: if α and β are reparametrizations by arc length of the same curve, then by Lemma 1.3.1 we have that $\alpha(s) = \beta(\pm s + c)$, so $\alpha'(s) = \pm\beta'(\pm s + c)$ and $\alpha''(s) = \beta''(\pm s + c)$. Therefore, $\|\alpha''(s)\| = \|\beta''(\pm s + c)\|$ and κ will be equal at each point of the curve independently of whether it is defined using α or β.

Example 1.4.1 Let us find the curvature of the curve $\alpha(t) = (e^t \cos t, e^t \sin t)$ of example 3.1. The parametrization by arc length of α is $\tilde{\alpha} = (\sqrt{s}\cos(\log\sqrt{s}), \sqrt{s}\sin(\log\sqrt{s}))$. Hence

$$\begin{aligned} \kappa(s) &= \|\tilde{\alpha}''(s)\| = \left\| \left(-\frac{\cos(\log(s/\sqrt{2})) + \sin(\log(s/\sqrt{2}))}{\sqrt{2}\, s}, \frac{\cos(\log(s/\sqrt{2})) - \sin(\log(s/\sqrt{2}))}{\sqrt{2}\, s} \right) \right\| \\ &= \frac{1}{s} \end{aligned}$$

Example 1.4.2 Let us find the curvature of a regular curve in $\mathbb{R}^2$ lying in a circumference of radius R centred at a point $\boldsymbol{p}$.

Let α be the curve and assume that it is parametrised by arc length. Since it lies in the circumference of radius R centred at a point $\boldsymbol{p}$, we have

$$\|\alpha(s) - \boldsymbol{p}\| = R.$$

We write the last equation as $(\alpha(s) - \boldsymbol{p}) \cdot (\alpha(s) - \boldsymbol{p}) = R^2$ and we take derivatives in both sides (**remember: if you do not know what to do, take derivatives!**). We obtain

$$2\alpha'(s) \cdot (\alpha(s) - \boldsymbol{p}) = 0,$$

so $\alpha'(s)$ and $(\alpha(s) - \boldsymbol{p})$ are perpendicular. Differentiating again,

$$2\alpha''(s) \cdot (\alpha(s) - \boldsymbol{p}) + 2\alpha'(s) \cdot \alpha'(s) = 0 \quad \text{so} \quad \alpha''(s) \cdot (\alpha(s) - \boldsymbol{p}) = -\alpha'(s) \cdot \alpha'(s).$$

On the other hand, α is parametrised by arc length so

$$\|\alpha'(s)\|^2 = \alpha'(s) \cdot \alpha'(s) = 1.$$

Differentiating the last equation we obtain

$$2\alpha'(s) \cdot \alpha''(s) = 0,$$

and therefore $\alpha''(s)$ must also be perpendicular to $\alpha'(s)$. Since we are in $\mathbb{R}^2$, two vectors perpendicular to a given one must be parallel, so $(\alpha(s) - \boldsymbol{p})$ and $\alpha''(s)$ must be parallel. Thus we have

$$|\alpha''(s) \cdot (\alpha(s) - \boldsymbol{p})| = \|\alpha''(s)\|\,\|(\alpha(s) - \boldsymbol{p})\| = \|\alpha'(s) \cdot \alpha'(s)\| = |-1| = 1.$$

Since $\|(\alpha(s) - \boldsymbol{p})\| = R$, we finally obtain

$$k(s) = \|\alpha''(s)\| = \frac{1}{R}.$$

1.5 Frenet frames, curvature and torsion

If we are riding in a roller coaster all the forces we feel are taken from the point of view of our reference system, which is riding with us in the roller coaster. So it makes sense to describe relevant quantities in a curve with a reference frame that is moving with the curve.

Let α be a parametrised regular curve. We define $\boldsymbol{T}$ as the unit tangent vector to α pointing towards the direction of movement of α. Hence, if α is parametrised by arc length,

$$\boldsymbol{T}(s) = \alpha'(s).$$

Although $\boldsymbol{T}$ is a function of the parameter s, we normally think of it as a function on the points $\alpha(s)$ of the curve. In fact...

IMPORTANT NOTE ABOUT 'ADDRESSES': all the quantities or vectors we define over the curves (or over the surfaces in the future) will be considered as defined over the curve itself (i.e. a function on the points that form the curve). However, each of this points can be addressed, via a one-to-one parametrization, by a value of the parameter. So although we could use the rather cumbersome notation $\boldsymbol{T}(\alpha(s))$, we write the equivalent shorter notation $\boldsymbol{T}(s)$.

When α is not parametrised by arc length, we define $\boldsymbol{T}(t) = \alpha'(t)/\|\alpha'(t)\|$.

Now note the following: if α is parametrised by arc length, then

$$\alpha'(s) \cdot \alpha'(s) = \|\alpha'(s)\|^2 = 1.$$

Taking derivatives in both sides we obtain

$$2\alpha'(s) \cdot \alpha''(s) = 0.$$

This implies that either $\alpha''(s) = 0$ or is perpendicular to $\alpha'(s)$.

In order to define our moving frame of reference we need three vectors. We already have $\boldsymbol{T}$. A good choice for another vector would be the unit vector pointing in the direction of $\alpha''(s)$. However we can only do this if $\alpha''(s) \neq 0$. So let us restrict our study to curves that satisfy this condition.

Definition 1.5.1 *A parametrised regular curve $\alpha : (a, b) \to \mathbb{R}^n$, parametrised by arc length, is called* biregular *if $\alpha''(s) \neq 0$ for all $s \in (a, b)$.*

A parametrised regular curve $\alpha : (a, b) \to \mathbb{R}^n$, NOT necessarily parametrised by arc length, is called biregular *if it has a reparametrization by arc length that is biregular (and hence all its reparametrizations by arc length will also be biregular—see the remark before example 4.1).*

Lemma 1.5.1 *A curve $\alpha : (a, b) \to \mathbb{R}^n$ is biregular if and only if $\alpha'(t)$ and $\alpha''(t)$ are linearly independent for all t.*

Proof : Let $\tilde{\alpha}(s) = \alpha(\phi(s))$ be a reparametrization by arc length of α. Then

$$\tilde{\alpha}'(s) = \alpha'(\phi(s))\, \phi'(s),$$

and

$$\tilde{\alpha}''(s) = \alpha''(\phi(s))\, (\phi'(s))^2 + \alpha'(\phi(s))\, \phi''(s).$$

If α is biregular then $\tilde{\alpha}'(s) \cdot \tilde{\alpha}''(s) = 0$ and $\tilde{\alpha}''(s) \neq 0$, so $\tilde{\alpha}'(s)$ and $\tilde{\alpha}''(s)$ must be linearly independent. But looking at the expressions above this implies that $\alpha'(\phi(s))$ and $\alpha''(\phi(s))$ must also be linearly independent; otherwise $\alpha''(\phi(s))$ would be a multiple of $\alpha'(\phi(s))$ so both $\tilde{\alpha}''(s)$ and $\tilde{\alpha}'(s)$ would be multiples of $\alpha'(\phi(s))$, contradicting their linear independence.

Conversely, if $\alpha'(t)$ and $\alpha''(t)$ are linearly independent then $\tilde{\alpha}''(s)$ cannot be 0 since it is a linear combination of $\alpha'(t)$ and $\alpha''(t)$ with nonzero coefficients (since $\phi'(s)$ is never 0).

□

Given a parametrised biregular curve, parametrised by arc length, we define

$$\boldsymbol{N}(s) = \frac{\alpha''(s)}{\|\alpha''(s)\|}.$$

Note that $\boldsymbol{N}$ has length 1 and is perpendicular to $\boldsymbol{T}$. Also, by definition of κ, $\boldsymbol{T}$ and $\boldsymbol{N}$, we have the formula

$$\boldsymbol{T}' = \kappa \boldsymbol{N}.$$

Up to this point everything we have done is for curves in $\mathbb{R}^n$, without specifying the values of n. From now on we have to make a distinction. First, in this course we will only study curves in $\mathbb{R}^2$ and $\mathbb{R}^3$. The same ideas work, however, for any value of n.

1.5.1 Curves in $\mathbb{R}^2$

We have defined the perpendicular vectors $\boldsymbol{T}$ and $\boldsymbol{N}$ that move with the curve. Therefore we already have a frame of reference for $\mathbb{R}^2$, given, for each value of the parameter s, by $\{\boldsymbol{T}(s), \boldsymbol{N}(s)\}$. We already know that $\boldsymbol{T}'(s) = \kappa(s)\boldsymbol{N}(s)$. What is the derivative of $\boldsymbol{N}$?

Since $\boldsymbol{N}$ has length 1, using the same calculation we did for α', we have that $\boldsymbol{N}'$ must be perpendicular to $\boldsymbol{N}$, and therefore it must be a multiple of $\boldsymbol{T}$. Now we use the following 'trick', which is important in many calculations about curves:

$$0 = (\boldsymbol{T} \cdot \boldsymbol{N})' = \boldsymbol{T}' \cdot \boldsymbol{N} + \boldsymbol{T} \cdot \boldsymbol{N}'.$$

Since $\boldsymbol{T}' = \kappa \boldsymbol{N}$, we have

$$\boldsymbol{T} \cdot \boldsymbol{N}' = -\kappa\,(\boldsymbol{N} \cdot \boldsymbol{N}) = -\kappa,$$

and therefore

$$\boldsymbol{N}' = -\kappa \boldsymbol{T}.$$

This gives the *Frenet-Serret* formulas for curves in $\mathbb{R}^2$:

$$\begin{cases} \boldsymbol{T}' &= \kappa \boldsymbol{N} \\ \boldsymbol{N}' &= -\kappa \boldsymbol{T} \end{cases}$$

If we think of the vectors $\boldsymbol{T}$ and $\boldsymbol{N}$ as column vectors in $\mathbb{R}^2$, written in some basis, then these equations can be written in matrix form as

$$(\boldsymbol{T}, \boldsymbol{N})' = (\boldsymbol{T}, \boldsymbol{N}) \begin{pmatrix} 0 & -\kappa \\ \kappa & 0 \end{pmatrix}.$$

It turns out that κ characterises completely the shape of biregular curves in $\mathbb{R}^2$. In other words, if we know κ, then we can describe the curve completely (although we will not know where it is located in the plane). But this is a consequence of a similar theorem in $\mathbb{R}^3$, so let us first study curves in $\mathbb{R}^3$.

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

A reference frame in space consists of three vectors. Given a biregular parametrised curve α we already defined the perpendicular unit vectors $\boldsymbol{T}$ and $\boldsymbol{N}$. A sensible choice for a third vector, so that it is unit and perpendicular to both $\boldsymbol{T}$ and $\boldsymbol{N}$ is

$$\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N},$$

where $\times$ is the cross product in $\mathbb{R}^3$.

The frame $\{\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B}\}$ is called the *Frenet frame* of α. Note that it is positively oriented, i.e. in the same way as the standard coordinate vectors in $\mathbb{R}^3$.

Let us calculate the derivative of $\boldsymbol{B}$. This calculation is a good example of how to calculate derivatives of vector fields over a curve. First, let us express $\boldsymbol{B}'$ in the frame $\{\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B}\}$:

$$\boldsymbol{B}' = (\boldsymbol{B}' \cdot \boldsymbol{T})\, \boldsymbol{T} + (\boldsymbol{B}' \cdot \boldsymbol{N})\, \boldsymbol{N} + (\boldsymbol{B}' \cdot \boldsymbol{B})\, \boldsymbol{B}.$$

As before, since $\boldsymbol{B} \cdot \boldsymbol{B} = 1$, differentiating both sides we find that $\boldsymbol{B}' \cdot \boldsymbol{B} = 0$.

On the other hand, $\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N}$, so differentiating both sides we obtain

$$\boldsymbol{B}' = \boldsymbol{T}' \times \boldsymbol{N} + \boldsymbol{T} \times \boldsymbol{N}' = \kappa \boldsymbol{N} \times \boldsymbol{N} + \boldsymbol{T} \times \boldsymbol{N}' = \boldsymbol{T} \times \boldsymbol{N}'.$$

Hence $\boldsymbol{B}'$ is perpendicular to $\boldsymbol{T}$, and therefore $\boldsymbol{B}' \cdot \boldsymbol{T} = 0$. Therefore we have that $\boldsymbol{B}'$ is a multiple of $\boldsymbol{N}$, so we write

$$\boldsymbol{B}' = -\tau \boldsymbol{N},$$

where $\tau = -\boldsymbol{B}' \cdot \boldsymbol{N}$. There is no clear reason for the minus sign, it is just custom. The function τ is called the *torsion*.

We already know $\boldsymbol{T}'$ and $\boldsymbol{B}'$. Let us now find $\boldsymbol{N}'$, using the same strategy as before. We write

$$\boldsymbol{N}' = (\boldsymbol{N}' \cdot \boldsymbol{T})\, \boldsymbol{T} + (\boldsymbol{N}' \cdot \boldsymbol{N})\, \boldsymbol{N} + (\boldsymbol{N}' \cdot \boldsymbol{B})\, \boldsymbol{B}.$$

As before, $\boldsymbol{N}' \cdot \boldsymbol{N} = 0$ since $\boldsymbol{N}$ has length 1. To find $\boldsymbol{T} \cdot \boldsymbol{N}'$ we do

$$0 = (\boldsymbol{T} \cdot \boldsymbol{N})' = \boldsymbol{T}' \cdot \boldsymbol{N} + \boldsymbol{T} \cdot \boldsymbol{N}' = \kappa \boldsymbol{N} \cdot \boldsymbol{N} + \boldsymbol{T} \cdot \boldsymbol{N}' = \kappa + \boldsymbol{T} \cdot \boldsymbol{N}'.$$

Therefore we have $\boldsymbol{T} \cdot \boldsymbol{N}' = -\kappa$.

Doing the same with $\boldsymbol{B}$ instead of $\boldsymbol{T}$ we find

$$0 = (\boldsymbol{B} \cdot \boldsymbol{N})' = \boldsymbol{B}' \cdot \boldsymbol{N} + \boldsymbol{B} \cdot \boldsymbol{N}' = -\tau \boldsymbol{N} \cdot \boldsymbol{N} + \boldsymbol{B} \cdot \boldsymbol{N}' = -\tau + \boldsymbol{B} \cdot \boldsymbol{N}'.$$

Therefore we have $\boldsymbol{B} \cdot \boldsymbol{N}' = \tau$.

We can summarise these results in the following

Theorem 1.5.1 *(Frenet-Serret equations) Let α be a biregular curve in $\mathbb{R}^3$, parametrised by arc length. Let $\boldsymbol{T} = \alpha'$, $\kappa = \|\boldsymbol{T}'\|$, $\boldsymbol{N} = \boldsymbol{T}'/\kappa$, $\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N}$ and $\tau = \boldsymbol{B} \cdot \boldsymbol{N}'$. Then we have*

$$\begin{array}{rcrcr} \boldsymbol{T}' & = & & \kappa \boldsymbol{N} & \\ \boldsymbol{N}' & = & -\kappa \boldsymbol{T} & & +\tau \boldsymbol{B} \\ \boldsymbol{B}' & = & & -\tau \boldsymbol{N} & \end{array}$$

If we think of the vectors $\boldsymbol{T}$, $\boldsymbol{N}$ and $\boldsymbol{B}$ as column vectors in $\mathbb{R}^3$ written in some basis then these equations can be written in matrix form as

$$(\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B})' = (\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B}) \begin{pmatrix} 0 & -\kappa & 0 \\ \kappa & 0 & -\tau \\ 0 & \tau & 0 \end{pmatrix}.$$

Remark 1.5.1 The Frenet-Serret equations for curves in the plane are a particular case of those in space, as follows.

Given a plane curve in $\mathbb{R}^2$ we can *embed* this $\mathbb{R}^2$ in $\mathbb{R}^3$ as the xy-plane to obtain a curve in space. The vectors $\boldsymbol{T}$ and $\boldsymbol{N}$ will then lie in the xy-plane. Since $\boldsymbol{B}$ is perpendicular to both $\boldsymbol{T}$ and $\boldsymbol{N}$, it must be perpendicular to the xy-plane. Since it is also unitary, it has to be either $(0,0,1)$ or $(0,0,-1)$. But then we have

$$\boldsymbol{B}' = (0,0,\pm 1)' = 0,$$

which implies that the torsion, τ, is 0, so the Frenet-Serret equations read

$$\begin{cases} \boldsymbol{T}' & = \kappa \boldsymbol{N} \\ \boldsymbol{N}' & = -\kappa \boldsymbol{T} \\ \boldsymbol{B}' & = 0, \end{cases}$$

which say exactly the same about $\boldsymbol{T}$ and $\boldsymbol{N}$ as the Frenet-Serret equations in $\mathbb{R}^2$.

This will imply that all local properties of curves in $\mathbb{R}^2$ can be deduced as a particular case of curves in $\mathbb{R}^2$.

Example 1.5.1

Let us find the Frenet-Serret frame, the curvature and the torsion for the curve $\alpha(t) = (\cos t, \sin t, t)$, $-\infty < t < \infty$ (**a helix**). First we parametrise it by arc length:

$$s(t) = \int_0^t \sqrt{\sin^2 t + \cos^2 t + 1}\, dt = \sqrt{2}\, t.$$

Thus we have $s = \sqrt{2}\, t$, or equivalently $t = \sqrt{2}\, s$ (note that I am abusing notation here: I am using s and t both as functions and variables). Hence the curve

$$\left(\cos \frac{s}{\sqrt{2}}, \sin \frac{s}{\sqrt{2}}, \frac{s}{\sqrt{2}} \right)$$

is an arc length reparametrisation of α. We have

$$\boldsymbol{T} = \left(\frac{-\sin \frac{s}{\sqrt{2}}}{\sqrt{2}}, \frac{\cos \frac{s}{\sqrt{2}}}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right).$$

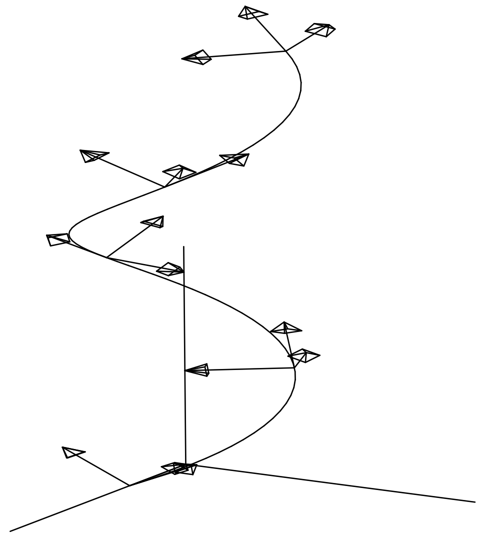
$$\kappa = \|\boldsymbol{T}'\| = \left\| \left(\frac{-\cos \frac{s}{\sqrt{2}}}{2}, \frac{-\sin \frac{s}{\sqrt{2}}}{2}, 0 \right) \right\| = \frac{1}{2}.$$

$$\boldsymbol{N} = \frac{\boldsymbol{T}'}{\kappa} = \left(-\cos \frac{s}{\sqrt{2}}, -\sin \frac{s}{\sqrt{2}}, 0 \right).$$

$$\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N} = \left(\frac{\sin \frac{s}{\sqrt{2}}}{\sqrt{2}}, \frac{-\cos \frac{s}{\sqrt{2}}}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$$

$$\tau = -\boldsymbol{B}' \cdot \boldsymbol{N} = -\left(\frac{\cos \frac{s}{\sqrt{2}}}{2}, \frac{\sin \frac{s}{\sqrt{2}}}{2}, 0 \right) \cdot \left(-\cos \frac{s}{\sqrt{2}}, -\sin \frac{s}{\sqrt{2}}, 0 \right) = \frac{1}{2}.$$

We obtain the following curve, where you can see the vectors $\boldsymbol{T}$, $\boldsymbol{N}$ and $\boldsymbol{B}$ at different points. Can you guess which is which?



1.5.3 Formulas for curves not parametrised by arc length

To find the curvature and torsion in the last example we had to first reparametrise the curve by arc length. This is annoying and sometimes extremely hard, since we have to integrate. The following are general formulas that are valid for any parametrisation of the curve.

Proposition 1.5.1 *Let $\alpha : (a, b) \to \mathbb{R}^3$ be a biregular parametrised curve. Let $Sp(t) = \|\alpha'(t)\|$ (the speed of α). Then*

$$\alpha' = Sp\,\boldsymbol{T} \qquad \alpha'' = Sp'\,\boldsymbol{T} + Sp^2\,\kappa\,\boldsymbol{N}$$

$$\alpha''' = (Sp'' - \kappa^2\,Sp^3)\ \boldsymbol{T} + (3\,Sp\ Sp'\,\kappa + Sp^2\,\kappa')\ \boldsymbol{N} + Sp^3\,\kappa\,\tau\,\boldsymbol{B},$$

where all the derivatives are taken with respect the parameter t.

Proof : Let $\tilde{\alpha}(s)$ be a reparametrization by arc length of α so that $\alpha(t) = \tilde{\alpha}(s(t))$, where $s(t)$ is the arc length function of α. We have $s'(t) = \|\alpha'(t)\| = Sp(t)$. We take derivatives carefully applying the chain rule and patiently simplify:

$$\alpha'(t) = \tilde{\alpha}'(s(t))\,s'(t) = Sp(t)\,\boldsymbol{T}(s(t)).$$

$$\alpha''(t) = [Sp(t)]^2\,\boldsymbol{T}'(s(t)) + Sp'(t)\,\boldsymbol{T}(s(t)) = Sp'(t)\,\boldsymbol{T}(s(t)) + [Sp(t)]^2\,\kappa(t)\,\boldsymbol{N}(s(t)).$$

$$\begin{aligned}
\alpha'''(t) &= Sp''(t)\,\boldsymbol{T}(s(t)) + Sp(t)\,Sp'(t)\,\boldsymbol{T}'(s(t)) + 2Sp(t)\,Sp'(t)\,\kappa(t)\,\boldsymbol{N}(s(t)) \\
&\qquad + [Sp(t)]^2\,\kappa'(t)\,\boldsymbol{N}(s(t)) + [Sp(t)]^3\,\kappa(t)\,\boldsymbol{N}'(s(t)) \\
&= Sp''(t)\,\boldsymbol{T}(s(t)) + 3\,Sp(t)\,Sp'(t)\,\kappa(t)\,\boldsymbol{N}(s(t)) + [Sp(t)]^2\,\kappa'(t)\,\boldsymbol{N}(s(t)) \\
&\qquad + [Sp(t)]^3\,\kappa(t)\,\big(-\kappa(t)\,\boldsymbol{T}(s(t)) + \tau(t)\,\boldsymbol{B}(s(t))\big) \\
&= \big(Sp''(t) - [Sp(t)]^3\,[\kappa(t)]^2\big)\,\boldsymbol{T}(s(t)) + \big(3\,Sp(t)\,Sp'(t)\,\kappa(t) + [Sp(t)]^2\,\kappa'(t)\big)\,\boldsymbol{N}(s(t)) \\
&\qquad + [Sp(t)]^3\,\kappa(t)\,\tau(t)\,\boldsymbol{B}(s(t))
\end{aligned}$$

□

Proposition 1.5.2 *Let $\alpha : (a, b) \to \mathbb{R}^3$ be a biregular parametrised curve. Then*

$$\boldsymbol{T}(t) = \frac{\alpha'(t)}{Sp(t)} \quad \boldsymbol{B}(t) = \frac{\alpha'(t) \times \alpha''(t)}{\|\alpha'(t) \times \alpha''(t)\|} \quad \boldsymbol{N} = \boldsymbol{B} \times \boldsymbol{T} \quad \kappa = \frac{\|\alpha' \times \alpha''\|}{\|\alpha'\|^3} \quad \tau = \frac{(\alpha' \times \alpha'') \cdot \alpha'''}{\|\alpha' \times \alpha''\|^2}.$$

PROOF : The first equation is immediate from Propositon 1.5.1. Also from Propositon 1.5.1 we obtain

$$\alpha' \times \alpha'' = Sp^3\,\kappa^2\,\boldsymbol{T} \times \boldsymbol{N} = Sp^3\,\kappa^2\,\boldsymbol{B}.$$

Since $Sp^3\,\kappa^2 > 0$, $\|\alpha' \times \alpha''\| = Sp^3\,\kappa^2$ and we obtain the desired expression.

$\boldsymbol{N} = \boldsymbol{B} \times \boldsymbol{T}$ since by definition $\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N}$.

For the expression for κ, using Propositon 1.5.1 we have

$$\alpha' \times \alpha'' = Sp^3\,\kappa\,\boldsymbol{T} \times \boldsymbol{N}.$$

Taking norms we obtain

$$\kappa = \frac{\|\alpha' \times \alpha''\|}{Sp^3} = \frac{\|\alpha' \times \alpha''\|}{\|\alpha'\|^3}.$$

The expression for τ comes from calculating $(\alpha' \times \alpha'') \cdot \alpha'''$ and using Proposition 1.5.1:

$$\begin{aligned}(\alpha' \times \alpha'') \cdot \alpha''' &= Sp^3\,\kappa\,(\boldsymbol{T} \times \boldsymbol{N}) \cdot \alpha''' \\ &= Sp^6\,\kappa^2\,\tau\,(\boldsymbol{B} \cdot \boldsymbol{B}) \quad \text{(the other terms dissapear)} \\ &= \|\alpha'\|^6 \left(\frac{\|\alpha' \times \alpha''\|}{\|\alpha'\|^3}\right)^2 \tau,\end{aligned}$$

from where we get the desired expression.

□

1.5.4 What are curvature and torsion good for?

We still have not justified our interest in the curvature and torsion. It turns out that these two functions tell us everything about the shape of the curve. Before proving this fact, we will prove a few results that are enlightening for understanding what curvature and torsion mean geometrically.

The following proposition is intuitively clear: since the curvature of a plane curve tells us how much it curves, if the curvature is constant it will always curve the same, so it will lie in a circumference.

PROPOSITION 1.5.3 *Let $\alpha : (a, b) \to \mathbb{R}^2$ be a biregular parametrised curve. Then the curvature of α is a constant $k > 0$ if and only if the image of α is contained in a circumference of radius $1/k$.*

PROOF :

We already know the ‘if’ part from Example 4.2.

For the ‘only if’ part, suppose that α has constant curvature k, and assume that it is parametrised by arc length (we know that there is no loss of generality on assuming this). Let us use some geometric intuition towards what we want to prove. If it were true that α is contained in a circumference of radius $1/k$, this could only be possible if there is a point $\boldsymbol{p}$ (the centre of this circumference) such that (draw a picture)

$$\alpha(s) + \frac{1}{k}\boldsymbol{N}(s) = \boldsymbol{p} \quad \text{(this is what we want to prove!)}$$

So we need to show that the curve $\alpha(s) + \frac{1}{k}\boldsymbol{N}(s)$ is constant. To this end, we take derivatives and use the Frenet-Serret equations:

$$(\alpha(s) + \frac{1}{k}\boldsymbol{N}(s))' = \alpha'(s) + \frac{1}{k}\boldsymbol{N}'(s)$$

$$\begin{aligned}&= \boldsymbol{T}(s) + \frac{1}{k}(-k\boldsymbol{T}(s)) \\ &= \boldsymbol{T}(s) - \boldsymbol{T}(s) \\ &= 0\end{aligned}$$

Thus the function $\alpha(s) + \frac{1}{k}\boldsymbol{N}(s)$ must be a constant which we will denote by $\boldsymbol{p}$, and we have

$$\alpha(s) + \frac{1}{k}\boldsymbol{N}(s) = \boldsymbol{p},$$

as desired. Reordering and taking norms we obtain

$$\|\alpha(s) - \boldsymbol{p}\| = \left\|\frac{1}{k}\boldsymbol{N}(s)\right\| = \frac{1}{k},$$

which implies that the curve lies in a circumference centred at $\boldsymbol{p}$ with radius $1/k$.

In fact it is an exercise in differential equations to show that $\alpha(s)$ must be of the form $\boldsymbol{p} + (\cos(\omega s + c), \sin(\omega s + c))$ for some constants ω and c. We will not do it here.

□

It would also make sense now that a curve with zero curvature would lie in a straight line (say, lies in a circle with infinite radius, i.e. a line). This is also true for curves in space even though the previous proposition is not (look at the helix of Example 5.1).

Proposition 1.5.4 *A regular curve with zero curvature lies in a straight line.*

Proof : Since the curvature is 0, we have $\alpha''(s) = 0$. Integrating twice we obtain $\alpha(s) = \boldsymbol{p} + s\vec{v}$ for some point $\boldsymbol{p}$ and some vector $\vec{v}$.

□

The following result gives us an interpretation of the torsion.

Proposition 1.5.5 *Let α be a biregular curve in space. Then the torsion of α is zero everywhere if and only if α lies in a plane.*

Proof : If α lies in a fixed plane, then $\boldsymbol{T}$ and $\boldsymbol{N}$ must be parallel to this plane, so $\boldsymbol{B}$ must be perpendicular to this fixed plane, so it does not change direction. Since in addition $\boldsymbol{B}$ has constant length equal to 1, it must be constant. Hence

$$0 = \boldsymbol{B}' = -\tau\boldsymbol{N},$$

which implies $\tau = 0$.

Conversely, suppose that $\tau = 0$. Then the last equation implies that $\boldsymbol{B}$ must be constant.

Let us use some geometric intuition again to see how we can express 'α lies in a plane' with formulas. The curve α will lie in a plane that passes through some fixed point (say $\boldsymbol{p}$) and is perpendicular to some vector (say $\vec{v}$) if and only if, for all s,

$$(\alpha(s) - \boldsymbol{p}) \cdot \vec{v} = 0 \quad \text{(this is what we want to prove!)}$$

This is the same as saying that the function $\alpha(s) \cdot \vec{v}$ is constant. A candidate for the vector $\vec{v}$ in our case would be the constant vector $\boldsymbol{B}$. Let us calculate

$$(\alpha(s) \cdot \boldsymbol{B}(s))' = \alpha'(s) \cdot \boldsymbol{B}(s) = \boldsymbol{T}(s) \cdot \boldsymbol{B}(s) = 0.$$

Thus $(\alpha(s) \cdot \boldsymbol{B}(s))$ is constant. If s_0 is some fixed value of s in the domain of $\alpha(s)$, we have

$$\alpha(s) \cdot \boldsymbol{B}(s) = \alpha(s_0) \cdot \boldsymbol{B}(s_0)$$

for all s, and therefore

$$(\alpha(s) - \alpha(s_0)) \cdot \boldsymbol{B}(s) = 0,$$

which expresses the fact that α lies in the plane passing through $\alpha(s_0)$ that is perpendicular to the (constant) vector $\boldsymbol{B}(s)$.

□

This proposition gives a geometrical interpretation of the torsion: it is a measure of how much a curve fails to be in a plane. We give now a more physical interpretation.

Suppose that you are riding in a roller coaster so that your head is always pointing in the direction of the normal $\boldsymbol{N}$ of the curve described by the roller coaster (this is the safest choice: the centrifugal force will only push us up or down, not to the sides, so it is more difficult to fall). If you stretch your arms in a cross then one arm will go up and the other will go down. The speed at which they move is the torsion.

And a most enlightening view of curvature and torsion is the following.

The local form of a curve

Let α be a biregular curve parametrised by arc length. With the help of Proposition 1.5.1 we can find the Taylor expansion of α up to order 3 about a point s_0. For convenience let us suppose that $s_0 = 0$ (otherwise reparametrise the curve). Since α is parametrised by arc length, $Sp(s) = 1$, so we have

$$\alpha'(0) = \boldsymbol{T}(0) \qquad \alpha''(0) = \kappa(0)\boldsymbol{N}(0)$$
$$\alpha'''(0) = -[\kappa(0)]^2\, \boldsymbol{T}(0) + \kappa'(0)\, \boldsymbol{N}(0) + \kappa(0)\, \tau(0)\, \boldsymbol{B}(0).$$

Thus, the Taylor expansion of α about 0 up to order 3 is

$$\begin{aligned} \alpha'(s) &= \alpha(0) + s\,\boldsymbol{T}(0) + s^2\,\frac{\kappa(0)\,\boldsymbol{N}(0)}{2} + s^3\,\frac{-[\kappa(0)]^2\,\boldsymbol{T}(0) + \kappa'(0)\,\boldsymbol{N}(0) + \kappa(0)\,\tau(0)\,\boldsymbol{B}(0)}{6} + \cdots. \\ &= \alpha(0) + \left(s - s^3\,\frac{[\kappa(0)]^2}{6} + \cdots\right)\,\boldsymbol{T}(0) \\ &\qquad + \left(s^2\,\frac{\kappa(0)}{2} + s^3\frac{\kappa'(0)}{6} + \cdots\right)\,\boldsymbol{N}(0) + \left(s^3\,\frac{\kappa(0)\,\tau(0)}{6} + \cdots\right)\,\boldsymbol{B}(0) \end{aligned}$$

Thus near the point $\alpha(0)$ the curve $\alpha(s)$ is approximately equal to the curve

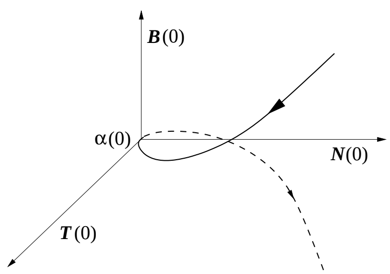
$$\beta(s) = \alpha(0) + \left(s - s^3\,\frac{[\kappa(0)]^2}{6}\right)\,\boldsymbol{T}(0) + \left(s^2\,\frac{\kappa(0)}{2} + s^3\frac{\kappa'(0)}{6}\right)\,\boldsymbol{N}(0) + \left(s^3\,\frac{\kappa(0)\,\tau(0)}{6}\right)\,\boldsymbol{B}(0).$$

The figures below represent the approximate graph of the curve $\beta(s)$ for some values of $\kappa(0)$ and $\tau(0)$. We use the axes with direction vectors $\boldsymbol{T}(0)$, $\boldsymbol{N}(0)$ and $\boldsymbol{B}(0)$ centred at $\alpha(0)$. The projections on the three coordinate planes are also depicted.

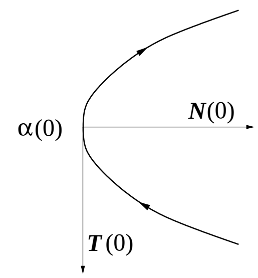
In general, the plane spanned by the vectors $\boldsymbol{T}$ and $\boldsymbol{N}$ is called the *osculating* plane (from the latin 'osculo', meaning 'kiss'). It is the plane that better adapts to the plane at that point. The plane spanned by $\boldsymbol{T}$ and $\boldsymbol{B}$ is called the *rectifying* plane, and the plane spanned by $\boldsymbol{N}$ and $\boldsymbol{B}$ is called the *normal* plane.

Observe the following:

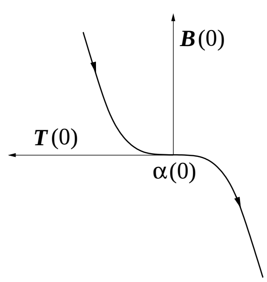
- The greater $\kappa(0)$ is, the narrower the 'parabola' is in the projection on the osculating plane, and the narrower the cubic is in the projection on the rectifying plane.
- If $\tau(0)$ is positive the projection on the rectifying and normal planes are as below. If $\tau(0)$ is negative the projection on the rectifying plane is a reflection on the $\boldsymbol{T}(0)$ axes of the one below, and the projection on the normal plane will be the same with the directions reversed.
- Note that if $\tau(0)$ is positive then the curve crosses the osculating plane at $\alpha(0)$ from the side at which $\boldsymbol{N}(0)$ points to the opposite side, whereas if $\tau(0)$ is negative it does so the other way around.
- The greater $\tau(0)$ is, the steeper will the two branches of the projection on the normal plane be.



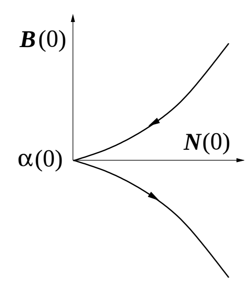
The curve $\beta(s)$



The projection on the $\boldsymbol{T}$, $\boldsymbol{N}$ (osculating) plane.



The projection on the $\boldsymbol{T}$, $\boldsymbol{B}$ (rectifying) plane



The projection on the $\boldsymbol{N}$, $\boldsymbol{B}$ (normal) plane.

1.5.5 Curvature and torsion determine the shape of the curve

Recall that what we understand by shape are those properties of a curve disregarding the position of the curve in space. In other words, if we have two curves and when we move one of them via *rigid motions*, i.e. rotations and translations, we obtain the other, we say that they have the same shape. In this section we will see that two biregular curves with the same curvature and torsion differ by a rigid motion.

Recall that a rigid motion is a rotation followed by a translation. A translation is just adding a fixed vector. A rotation is a linear map whose matrix A in the standard basis of $\mathbb{R}^n$ satisfies $\det(A) = 1$ and $AA^t = A^tA = I$.

Theorem 1.5.2 *Given smooth functions $\kappa(s) > 0$ and $\tau(s)$, $a < s < b$, there exists a biregular parametrised curve $\alpha : (a, b) \to \mathbb{R}^3$ such that s is the arc length and such that its curvature and torsion are given by $\kappa(s)$ and $\tau(s)$ respectively.*

Moreover any other curve $\beta(s)$ satisfying the same conditions differs from $\alpha(s)$ by a rigid motion; that is, there exists an orthonormal map and a vector $\vec{v}$ such that rigid motion $M : \mathbb{R}^3 \to \mathbb{R}^3$ such that $\beta(s) = \vec{v} + A(\alpha(s))$ for all s.

The proof of this theorem needs some important results. First of all we need the following theorem from ODE that we state without proof.

THEOREM 1.5.3 *Let $A : (a, b) \to \mathcal{M}_{n\times n}(\mathbb{R})$ be a smooth function, where $\mathcal{M}_{n\times n}(\mathbb{R})$ denotes the set of n by n matrices with real entries. Let $t_0 \in (a, b)$ and $\boldsymbol{x}_0 \in \mathbb{R}^n$.*

Then there exists a unique *function $\boldsymbol{x} : (a, b) \to \mathbb{R}^n$ such that*

$$\boldsymbol{x}'(t) = A(t)\boldsymbol{x}(t) \quad \textit{and} \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0.$$

Using this theorem we state and prove a refined version of the first half of Theorem 1.5.2.

PROPOSITION 1.5.6 *Let $\kappa(s) > 0$ and $\tau(s)$, $a < s < b$, be smooth functions. Let $s_0 \in (a, b)$ be a fixed number, let $\boldsymbol{p} \in \mathbb{R}^3$ be a point in $\mathbb{R}^3$ and let $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, with $\vec{e}_3 = \vec{e}_1 \times \vec{e}_2$, be a positively oriented orthonormal frame.*

Then there exists a unique *biregular parametrised curve $\alpha : (a, b) \to \mathbb{R}^3$ such that α is parametrised by arc length, such that its curvature and torsion are given by $\kappa(s)$ and $\tau(s)$ respectively, and such that $\alpha(s_0) = \boldsymbol{p}$, $\boldsymbol{T}(s_0) = \vec{e}_1$, $\boldsymbol{N}(s_0) = \vec{e}_2$ and $\boldsymbol{B}(s_0) = \vec{e}_3$.*

PROOF :

Set the following system of differential equations (in $\mathbb{R}^{12}$):

$$\begin{cases} \alpha'(s) &= \boldsymbol{T}(s) \\ \boldsymbol{T}'(s) &= \kappa(s)\boldsymbol{N}(s) \\ \boldsymbol{N}'(s) &= -\kappa(s)\boldsymbol{T}(s) + \tau(s)\boldsymbol{B}(s) \\ \boldsymbol{B}'(s) &= -\tau(s)\boldsymbol{N}(s) \end{cases}$$

with the conditions

$$\begin{cases} \alpha(s_0) &= \boldsymbol{p} \\ \boldsymbol{T}(s_0) &= \vec{e}_1 \\ \boldsymbol{N}(s_0) &= \vec{e}_2 \\ \boldsymbol{B}(s_0) &= \vec{e}_3 \end{cases}$$

(These equations are just the definition of $\boldsymbol{T}$ and the Frenet-Serret equations.) It can be written in the form specified in the hypotheses of the last theorem, so this implies that there is a unique solution satisfying the given condition. It only remains to prove that the functions $\boldsymbol{T}$, $\boldsymbol{N}$ and $\boldsymbol{B}$ form the Frenet frame of the curve $\alpha(s)$. The proof of this fact is not trivial.

Suppose that $\alpha(s)$, $\boldsymbol{T}(s)$, $\boldsymbol{N}(s)$ and $\boldsymbol{B}(s)$ are a solution of the system of equations above. Then the dot products of the vectors $\boldsymbol{T}$, $\boldsymbol{N}$ and $\boldsymbol{B}$ must implicitly satisfy the following (this is a matter of applying the product rule in the derivatives of the dot products and using the expression for $\boldsymbol{T}'$, $\boldsymbol{N}'$ and $\boldsymbol{B}'$ above):

$$\begin{array}{rcl} (\boldsymbol{T}(s)\cdot\boldsymbol{T}(s))' &=& 2\kappa(s)\,(\boldsymbol{T}(s)\cdot\boldsymbol{N}(s)) \\ (\boldsymbol{T}(s)\cdot\boldsymbol{N}(s))' &=& \kappa(s)\,(\boldsymbol{N}(s)\cdot\boldsymbol{N}(s)) - \kappa(s)\,(\boldsymbol{T}(s)\cdot\boldsymbol{T}(s)) + \tau(s)\,(\boldsymbol{T}(s)\cdot\boldsymbol{B}(s)) \\ (\boldsymbol{T}(s)\cdot\boldsymbol{B}(s))' &=& \tau(s)\,(\boldsymbol{T}(s)\cdot\boldsymbol{N}(s)) - \kappa(s)\,(\boldsymbol{N}(s)\cdot\boldsymbol{B}(s)) \\ (\boldsymbol{N}(s)\cdot\boldsymbol{N}(s))' &=& -2\kappa(s)\,(\boldsymbol{N}(s)\cdot\boldsymbol{T}(s)) + 2\tau(s)\,(\boldsymbol{N}(s)\cdot\boldsymbol{B}(s)) \\ (\boldsymbol{N}(s)\cdot\boldsymbol{B}(s))' &=& -\kappa(s)\,(\boldsymbol{T}(s)\cdot\boldsymbol{B}(s)) - \tau(s)\,(\boldsymbol{N}(s)\cdot\boldsymbol{N}(s)) + \tau(s)\,(\boldsymbol{B}(s)\cdot\boldsymbol{B}(s)) \\ (\boldsymbol{N}(s)\cdot\boldsymbol{N}(s))' &=& -2\tau(s)\,(\boldsymbol{N}(s)\cdot\boldsymbol{B}(s)) \end{array}$$

with initial conditons

$$\begin{cases} \boldsymbol{T}(s_0)\cdot\boldsymbol{T}(s_0) &= 1 \\ \boldsymbol{N}(s_0)\cdot\boldsymbol{N}(s_0) &= 1 \\ \boldsymbol{B}(s_0)\cdot\boldsymbol{B}(s_0) &= 1 \\ \boldsymbol{T}(s_0)\cdot\boldsymbol{N}(s_0) &= 0 \\ \boldsymbol{T}(s_0)\cdot\boldsymbol{B}(s_0) &= 0 \\ \boldsymbol{N}(s_0)\cdot\boldsymbol{B}(s_0) &= 0 \end{cases}$$

Note that this is not a system on the vectors $\boldsymbol{T}$, $\boldsymbol{N}$ and $\boldsymbol{B}$ but on all their possible pairwise dot products. The previous theorem guarantees that the system has a unique solution. But note that

$$\begin{cases} \boldsymbol{T}(s)\cdot\boldsymbol{T}(s) & = 1 \\ \boldsymbol{N}(s)\cdot\boldsymbol{N}(s) & = 1 \\ \boldsymbol{B}(s)\cdot\boldsymbol{B}(s) & = 1 \\ \boldsymbol{T}(s)\cdot\boldsymbol{N}(s) & = 0 \\ \boldsymbol{T}(s)\cdot\boldsymbol{B}(s) & = 0 \\ \boldsymbol{N}(s)\cdot\boldsymbol{B}(s) & = 0 \end{cases}$$

is 'also' a solution (this needs to be checked - we leave it for the reader). Thus it must be *the* solution. Therefore the frame formed by the vectors $\boldsymbol{T}$, $\boldsymbol{N}$ and $\boldsymbol{B}$ is orthonormal. In addition, since $\boldsymbol{B}$ has length one and is perpendicular to both $\boldsymbol{T}$ and $\boldsymbol{N}$, it must be either equal to $\boldsymbol{T}\times\boldsymbol{N}$ or $-\boldsymbol{T}\times\boldsymbol{N}$. It cannot be one of them for some values of the parameter and the other one for other values because this would contradict continuity of $\boldsymbol{B}$. Since at s_0 we have $\boldsymbol{T}\times\boldsymbol{N}=\boldsymbol{B}$, we must thus have that this equation is true for all values of the parameter. Therefore the frame formed by $\boldsymbol{T}$, $\boldsymbol{N}$ and $\boldsymbol{B}$ must be the Frenet frame for $\alpha(s)$ since it satisfies $\boldsymbol{T}=\alpha'$, $\boldsymbol{N}=\boldsymbol{T}'/\kappa$ and $\boldsymbol{B}=\boldsymbol{T}\times\boldsymbol{N}$.

□

Now we are ready to prove the second part of Theorem 1.5.2.

PROOF : Suppose that $\alpha(s)$ and $\beta(s)$ have the same curvature $\kappa(s)$ and the same torsion $\tau(s)$ for every s. Let s_0 be a point in the domain of α and β, and let $\{\boldsymbol{T}_\alpha(s_0),\boldsymbol{N}_\alpha(s_0),\boldsymbol{B}_\alpha(s_0)\}$ and $\{\boldsymbol{T}_\beta(s_0),\boldsymbol{N}_\beta(s_0),\boldsymbol{B}_\beta(s_0)\}$ be the Frenet frames at the point s_0 of α and β, respectively. Let A be a 3 by 3 orthogonal matrix with determinant 1 that satisfies

$$\begin{aligned} A\,\boldsymbol{T}_\alpha(s_0) &= \boldsymbol{T}_\beta(s_0) \\ A\,\boldsymbol{N}_\alpha(s_0) &= \boldsymbol{N}_\beta(s_0) \\ A\,\boldsymbol{B}_\alpha(s_0) &= \boldsymbol{B}_\beta(s_0) \end{aligned}$$

(Note that both Frenet frames are orthonormal and positively oriented so such an orthogonal map must exist—check Chapter 0.) Let us calculate the curvature $\tilde{\kappa}(s)$ and torsion $\tilde{\tau}(s)$ of the curve $\tilde{\alpha}(s)=\vec{v}+A(\alpha(s))$, where $\vec{v}=\beta(s_0)-A(\alpha(s_0))$. First we have

$$\tilde{\alpha}'(s)=DA_{\alpha(s)}(\alpha'(s))=A(\alpha'(s)),$$

since the derivative of a linear function is itself, and also

$$\tilde{\alpha}''(s)=DA_{\alpha'(s)}(\alpha''(s))=A(\alpha''(s)),$$

$$\tilde{\alpha}'''(s)=DA_{\alpha''(s)}(\alpha'''(s))=A(\alpha'''(s)).$$

Let us use the formulas in section 5.3.

$$\begin{aligned} \tilde{\kappa}(s) &= \frac{\|\tilde{\alpha}'(s)\times\tilde{\alpha}''(s)\|}{\|\tilde{\alpha}'(s)\|^3} \\ &= \frac{\|A(\alpha'(s))\times A(\alpha''(s))\|}{\|A(\alpha'(s))\|^3} \\ &= \frac{\|\det(A)A(\alpha'(s)\times\alpha''(s))\|}{\|A(\alpha'(s))\|^3} \\ &= \frac{\|\alpha'(s)\times\alpha''(s)\|}{\|\alpha'(s)\|^3} \\ &= \kappa(s), \end{aligned}$$

since A preserves norms and has determinant 1 since it is orthonormal.

$$\begin{aligned}
\tilde{\tau}(s) &= \frac{(\tilde{\alpha}'(s) \times \tilde{\alpha}''(s)) \cdot \tilde{\alpha}'''(s)}{\|\tilde{\alpha}'(s) \times \tilde{\alpha}''(s)\|^2} \\
&= \frac{[A(\alpha'(s)) \times A(\alpha''(s))] \cdot A(\alpha'''(s))}{\|A(\alpha'(s)) \times A(\alpha''(s))\|^2} \\
&= \frac{\det(A)\ ((\alpha'(s) \times \alpha''(s)) \cdot \alpha'''(s))}{\|\det(A)\ A(\alpha'(s) \times \alpha''(s))\|^2} \\
&= \frac{(\alpha'(s) \times \alpha''(s)) \cdot \alpha'''(s)}{\|\alpha'(s) \times \alpha''(s)\|^2} \\
&= \tau(s)
\end{aligned}$$

To finish the proof, note that $\tilde{\alpha}(s)$ and $\beta(s)$ have the same curvature and torsion, and the same Frenet frame at the common point $\beta(s_0)$ (note that $\tilde{\alpha}(s_0) = \beta(s_0)$). Then, by the uniqueness part of the last proposition, they must be the same curve. Hence , writing $\boldsymbol{p} = \beta(s_0) - A(\alpha(s_0))$, we have

$$\beta(s) = \vec{v} + A(\alpha(s)) \quad \text{for all } s \in (a, b),$$

so α and β differ by a rigid motion, as claimed.

$\square$

Chapter 2

Surfaces

2.1 What is a surface?

2.1.1 Regular surfaces and parametrised surfaces

As in the case of curves, we make two definitions of the concept of surface. One of them (regular surface) emphasizes the fact that a surface, as we think of it, is a set of points. The other (parametrised surface) emphasizes the parametrization of the surface. While these two concepts were similar in the case of curves (every regular curve can be covered with a single parametrization, so it is a parametrised regular curve), they are different for surfaces: a sphere, for example, is a regular surface, but not a parametrised regular surface.

From now on we specialise for the case of surfaces in $\mathbb{R}^3$. For the definitions for surfaces in $\mathbb{R}^n$, any n, just substitute 3 for n.

Definition 2.1.1 *A* parametrised surface *is a smooth map $\sigma : U \subset \mathbb{R}^2 \to \mathbb{R}^3$, where U is open in $\mathbb{R}^2$. The surface is called* regular parametrised surface *if the map σ is regular.*

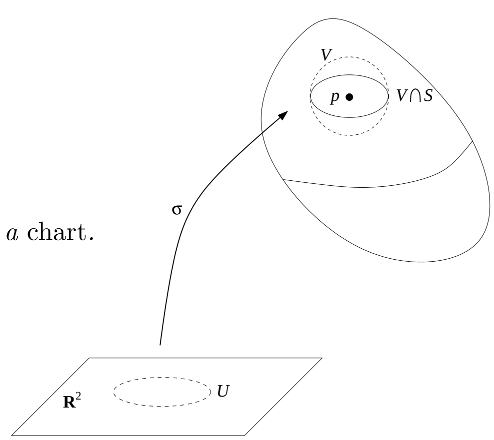
The regularity condition is there to guarantee that the image of σ is indeed what we intuitively understand as a surface (otherwise, taking σ constant—so not regular—we would have that a point is a surface, which is counterintuitive).

Definition 2.1.2 *A* regular surface *S is a subset of $\mathbb{R}^3$ such that for every point $p \in S$ there is an open set $V \in R^3$, an open set $U \in \mathbb{R}^2$ and a map $\sigma : U \to S \cap V \subset \mathbb{R}^3$ satisfying*

- *σ is a homeomorphism.*
- *σ is smooth.*
- *σ is regular.*

Each map $\sigma : U \to S \cap V \subset \mathbb{R}^3$ of this sort is called a surface patch *or* parametrization*; σ^{-1} is often called a* chart*.*

This second definition probably looks complicated.
Let us explain what it means. First let us do a picture.



The situation depicted above is *for each p in S*. In other words, a surface is some subset of $\mathbb{R}^3$ that can be covered by surface patches. Each surface patch looks like a (maybe deformed) piece of $\mathbb{R}^2$.

A set of surface patches covering S is called an *atlas*.

Think of the surface patches as if they were maps (in the cartographic sense) of parts of the surface, and an atlas as a collection of all these maps (like an 'atlas of the world' but with the surface).

Example 2.1.1

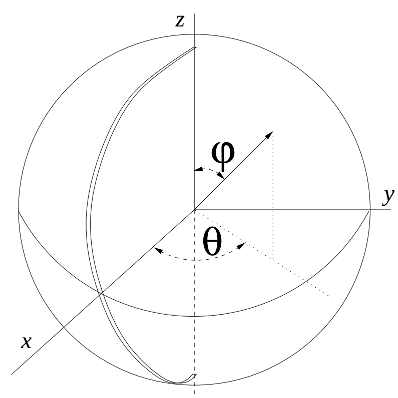
The unit sphere

$$S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$$

is a regular surface in $\mathbb{R}^3$. To see this we have to find functions σ from an open set in $\mathbb{R}^2$ to the sphere so that every point of the sphere is covered by a chart. There are many ways to do this. For example,

$$\sigma_1 : (0, 2\pi) \times (0, \pi) \to \mathbb{R}^3, \quad \sigma_1(\theta, \phi) = (\cos\theta\cos\phi, \sin\theta\cos\phi, \sin\phi).$$

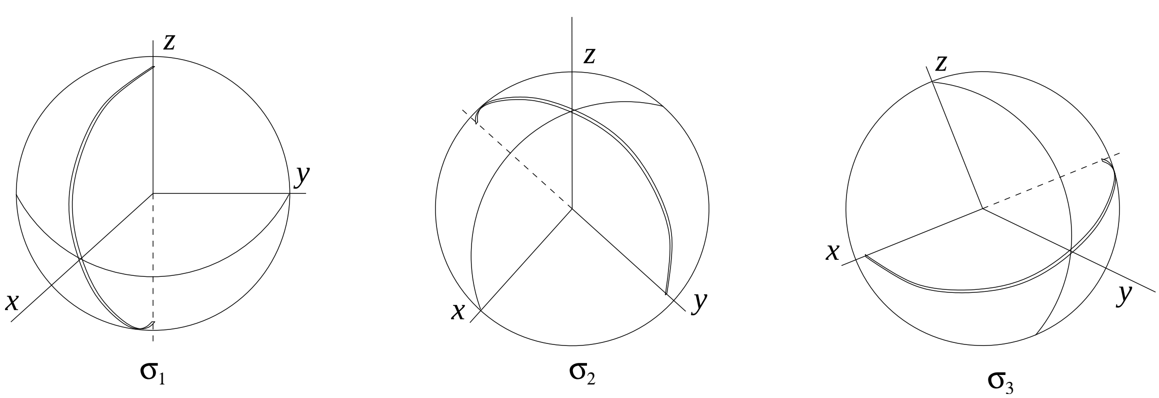
This covers all of the sphere except the slit formed by the points at $y = 0$, $x > 0$:



If we want to cover the rest of the points, we can just rotate our charts, i.e. rotate the coordinates in σ_1:

$$\begin{aligned}
\sigma_2 &: (0, 2\pi) \times (0, \pi) \to \mathbb{R}^3, & \sigma(\theta, \phi) &= (\sin\phi, \cos\theta\cos\phi, \sin\theta\cos\phi) \\
\sigma_3 &: (0, 2\pi) \times (0, \pi) \to \mathbb{R}^3, & \sigma(\theta, \phi) &= (\sin\theta\cos\phi, \sin\phi, \cos\theta\cos\phi)
\end{aligned}$$

The portions of the sphere covered by each patch are shown below. You can see that σ_1 and σ_2 cover everything but the point $(0, 0, 1)$, but this is covered by σ_3.



In the case of curves, all regular curves were parametrised regular curves, or equivalently they were covered by a single coordinate patch, this is not the case for surfaces. For example, a sphere *cannot* be covered by a single coordinate patch (think of wrapping a ball with paper; then when you try to close it many points of the wrapping will cover very few points of the sphere. At these points the coordinate patch will not be regular). On the other hand parametrised regular surfaces with a self intersection will not be regular surfaces (at the self intersection they do not look like a piece of the plane, as regular surfaces should).

So it may seem that we should treat these two cases separately. However, note that the portion of a regular surface parametrised by a single patch is a regular parametrised surface, so the study of local properties of surfaces is essentially the same for both cases.

In any case, in this course we will concentrate our study in regular surfaces, i.e. we will exclude surfaces with self-intersections.

2.1.2 How to see that something is a surface

It would be very annoying if we had to check all the properties in Definition 1.2 to see that something is a surface. Fortunately we have some simpler criteria.

Proposition 2.1.1 *If $f : U \in \mathbb{R}^2 \to \mathbb{R}$ is smooth, then the graph of f is a regular surface that is covered by a single patch.*

Proof :

Recall that the graph of a function f is the set of points of the form $\{(x, y, f(x, y)) : x, y \in U\}$.

So let $\sigma : U \to \mathbb{R}^3$ be defined as $\sigma(u, v) = (u, v, f(u, v))$ (we could use x and y as variables but for some reason it is customary to use u and v). Let us check that σ is an acceptable patch.

First by definition of the graph of f, σ covers all of it, so it is onto. It is also injective since different (u, v) will give different $(u, v, f(u, v))$, so σ is bijective.

σ is continuous and smooth because u, v and $f(u, v)$ are. The inverse of σ is just the projection $\pi(u, v, f(u, v)) = (u, v)$, which is also continuous.

Finally,

$$D\sigma_{(u,v)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ f_u(u, v) & f_v(u, v) \end{pmatrix},$$

which has rank 2, and therefore it is regular.

□

Another useful criterion is

Proposition 2.1.2 *Let $F : \mathbb{R}^3 \to \mathbb{R}$ be a smooth function. Consider the set $S = \{(x, y, z) : F(x, y, z) = 0\}$. If $\nabla F \neq 0$ at every point of S, then S is a regular surface.*

Proof : (sketch) This is a consequence of the implicit function theorem, that essentially says that given an equation of the form $F(x, y, z) = 0$ with F regular then one can always solve for one of the variables, at least locally. So to fix ideas suppose that we can solve for z, i.e. we can write

$$F(x, y, z) = 0 \iff z = f(x, y).$$

Then S will be locally the graph of the function f.

Note that we cannot expect all of S to be the graph of a function, as the following example shows.

□

Example 2.1.2

Consider $F(x,y,z) = x^2+y^2+z^2-1$. Then $\nabla f = (2x, 2y, 2z)$ which is not zero when $F(x,y,z) = 0$ (since if $x^2+y^2+z^2-1 = 0$ then either x, y or z is not 0). So the set of points where F is 0 is a regular surface—a sphere as you know—normally denoted as S^2.

Let us try to patch S^2 with graphs of functions:

- The graph of $z = \sqrt{1-x^2-y^2}$ covers the top open hemisphere.
- The graph of $z = -\sqrt{1-x^2-y^2}$ covers the bottom open hemisphere.
- The graph of $y = \sqrt{1-x^2-z^2}$ covers the right open hemisphere.
- The graph of $y = -\sqrt{1-x^2-z^2}$ covers the left open hemisphere.
- The graph of $x = \sqrt{1-y^2-z^2}$ covers the front open hemisphere.
- The graph of $x = -\sqrt{1-y^2-z^2}$ covers the back open hemisphere.

These six patches cover all of S^2. Note that neither of these patches can only be defined when whatever is inside the square root is greater than 0, and cannot be extended any further.

Example 2.1.3

Surfaces of revolution: consider the surface M obtained by rotating a curve $(f(v), 0, g(v))$, $v \in (a,b)$ lying in the xz-plane about the z-axes. To guarantee regularity of M we assume that the curve is regular and that $f(v) > 0$ for all v.

A patch for M will be

$$\sigma(u,v) = (f(v)\cos u, f(v)\sin u, g(v)), \quad 0 < u < 2\pi, a < v < b.$$

σ covers all of M except for the slit corresponding to the original curve $(f(v), 0, g(v))$. In order to cover this slit let

$$\tau(u,v) = (f(v)\cos u, f(v)\sin u, g(v)), \quad -\pi < u < \pi, a < v < b.$$

Then σ and τ together cover all of M.

Note that σ is regular (for τ is the same computation):

$$\sigma_u = (-f(v)\sin u, f(v)\cos u, 0) \text{ and } \sigma_v = (f'(v)\cos u, f'(v)\sin u, g'(v)).$$

$$\sigma_u \times \sigma_v = (f(v)g'(v)\cos u, f(v)g'(v)\sin u, -f(v)f'(v)),$$

which is never 0 since $f(v) > 0$ and $f'(v)$ or $g'(v)$ cannot both be 0 (the original curve is regular).

Of course we could have started with a curve in any vertical plane instead of the xz-plane; it works the same way. We could also have rotated a curve about any of the other axes, provided that the original curve is chosen accordingly.

2.2 The tangent plane

Definition 2.2.1 *Let M be a regular surface and $p \in M$. The* tangent plane of M at p *is*

$$T_pM = \{\alpha'(t_0) \text{ where } \alpha : (a,b) \to M \text{ smooth curve and } \alpha(t_0) = p\}$$

In other words, T_pM is the set of all possible tangent vectors to curves in M at the point p.

Lemma 2.2.1 *Let $\sigma : U \subset \mathbb{R}^2 \to M$ be a surface patch with $p = \sigma(u_0, v_0)$. Then*

$$T_pM = \operatorname{span}\{\sigma_u(u_0, v_0), \sigma_v(u_0, v_0)\}$$

Proof : First we show that $T_pM \supset \operatorname{span}\{\sigma_u(u_0, v_0), \sigma_v(u_0, v_0)\}$. Consider the curve

$$\alpha(t) = \sigma(u_0 + at, v_0 + bt),$$

where a and b are arbitrary numbers. Now, $\alpha(0) = \sigma(u_0, v_0) = p$, so $\alpha'(0) \in T_pM$. However, the chain rule gives

$$\alpha'(0) = \frac{\partial \sigma}{\partial u}(u_0, v_0)\, a + \frac{\partial \sigma}{\partial v}(u_0, v_0)\, b = a\sigma_u(u_0, v_0) + b\sigma_v(u_0, v_0),$$

proving that $T_pM \supset \operatorname{span}\{\sigma_u(u_0, v_0), \sigma_v(u_0, v_0)\}$.

Now we prove that $T_pM \subset \operatorname{span}\{\sigma_u(u_0, v_0), \sigma_v(u_0, v_0)\}$. If $\alpha : (a,b) \to M$ is a curve with $\alpha(t_0) = p$ then for t in some subinterval (a', b') the curve will lie in $\sigma(U)$, and hence we can write $\alpha(t) = \sigma(\beta(t))$ for some curve $\beta(\beta_1(t), \beta_2(t)) : (a', b') \to U \subset \mathbb{R}^2$. (We take as given the fact that β is smooth; this is a consequence of the inverse function theorem.) Then, using the chain rule,

$$\alpha'(t_0) = \beta_1'(t_0)\, \sigma_u(u_0, v_0) + \beta_2'(t_0)\, \sigma_v(u_0, v_0) \in \operatorname{span}\{\sigma_u(u_0, v_0), \sigma_v(u_0, v_0)\}.$$

$\square$

Remark 2.2.1

Since T_pM is the span of two vectors, it must be a vector subspace of $\mathbb{R}^3$.

Strictly speaking, T_pM should be defined as the space of pairs $(p, \alpha'(t_0)) \in \mathbb{R}^3 \times \mathbb{R}^3$, where $\alpha : (a,b) \to M$ is a curve such that $\alpha(t_0) = p$. The vector addition and multiplication by scalars in T_pM would be then defined as $\lambda(p, \vec{v}) + (p, \vec{w}) = (p, \lambda\vec{v} + \vec{w})$. We will not use this formalism to avoid complication with notation.

However, it is always convenient to think of T_pM as the vector space specified in the definition but 'anchored at the point p'. Geometrically we could think of it as the plane that is 'tangent' to M at p, but then this plane would not pass, in general, through the 0 of $\mathbb{R}^3$, so it would not be a vector space. So think of T_pM as a vector space, again, 'anchored' at p, or if you want, with a tag indicating that it is tangent at the point p.

Example 2.2.1 The tangent space to M parametrised by $\sigma(u, v) = (u, v, u^2 + v^2)$ at the point $p = (1, 1, 2) = \sigma(1, 1)$ is the subspace of $\mathbb{R}^3$ given by

$$\{a\, \sigma_u(1,1) + b\, \sigma_v(1,1)\} = \left\{ a \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right\}.$$

2.3 Functions on surfaces

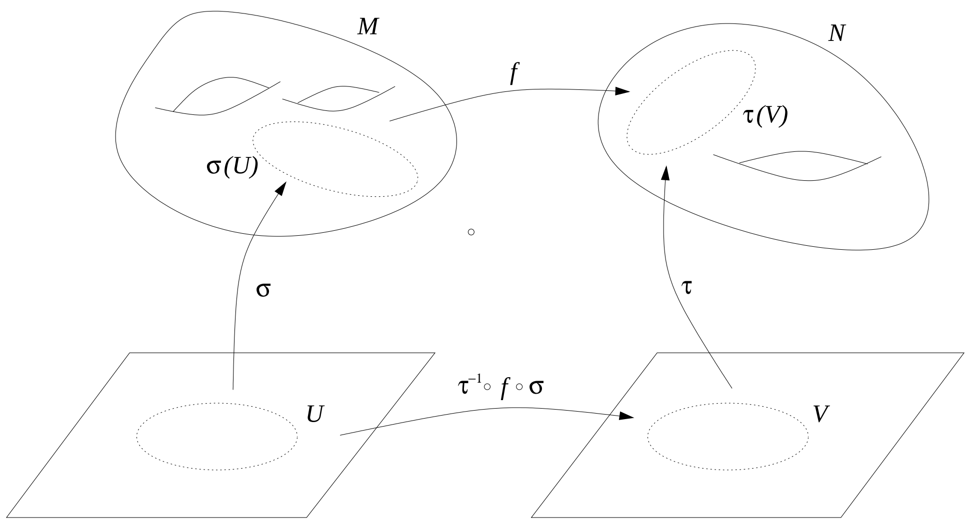
Definition 2.3.1 *Let M and N be a regular surfaces. A function $f : M \to \mathbb{R}^n$ is said to be differentiable if $f \circ \sigma : U \to \mathbb{R}^n$ is differentiable for every surface patch $\sigma : U \to M$ of M.*

A function $f : M \to N \subset \mathbb{R}^3$ is differentiable (as a function from M to N) if it is differentiable as a function from M to $\mathbb{R}^3$.

Given charts $\sigma : U \to M$ and $\tau : V \to N$ the expression

$$\tau^{-1} \circ f \circ \sigma : U \to V$$

(wherever it is defined) is called 'local expression of f in the patches σ to τ'. It turns out that f is differentiable if and only if the local expression in any patches is differentiable.



2.4 Vector fields

Definition 2.4.1 *Let $M \subset \mathbb{R}^3$ be a regular surface. A vector field in M is a map*

$$\begin{aligned} X : M &\to \mathbb{R}^3 \\ p &\to X_p \end{aligned}$$

where we think of the value of X at a point p as a vector in $\mathbb{R}^3$ 'anchored' at p (whatever this may mean...). To emphasize this fact we normally write X_p instead of $X(p)$ for the value of X at p (although we will use both notations).

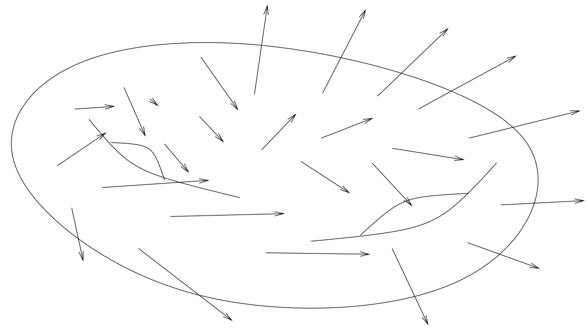
A vector field is tangent *if $X_p \in T_pM$ $\forall p$, and* normal *if $X_p \perp T_pM$ $\forall p$.*

Remark 2.4.1

Let $T\mathbb{R}^3$ be the set of pairs $(p, v) \subset \mathbb{R}^3 \times \mathbb{R}^3$ with addition and multiplication by scalars defined by $\lambda(p, v) + (p, w) = (p, \lambda v + w)$. Then, strictly speaking, a vector field is a map from M to $T\mathbb{R}^3$ that satisfies $X(p) = (p, X_p)$.

This view of vector fields is certainly more correct but for our present purposes would only complicate matters, so we will think of vector fields just as maps from M to $\mathbb{R}^3$.

Example 2.4.1



Example 2.4.2

Coordinate vector fields:

Let $\sigma(u, v)$ be a coordinate patch of M. Then $\sigma_u(u, v)$ and $\sigma_v(u, v)$ are tangent vectors at $\sigma(u, v)$, so they are tangent vector fields defined, in general, on part of the surface M (σ my not cover all of M). They are called 'coordinate vector fields'.

To emphasize the fact that they are vector fields we will write

$$(\sigma_u)_p \quad \text{and} \quad (\sigma_v)_p$$

instead of

$$\sigma_u(u_0, v_0) \quad \text{and} \quad \sigma_v(u_0, v_0),$$

where it is understood that $p = \sigma(u_0, v_0)$. In other words, we define

$$(\sigma_u)_{\sigma(u,v)} := \sigma_u(u, v) \quad \text{and} \quad (\sigma_v)_{\sigma(u,v)} := \sigma_v(u, v).$$

Each of these two notations has advantages and disadvantages, so we will jump from one notation to the other when convenient.

Also, as it is customary, we will often simply write

$$\sigma_u \quad \text{and} \quad \sigma_v$$

instead of the more cumbersome $\sigma_u(u, v)$ and $\sigma_v(u, v)$.

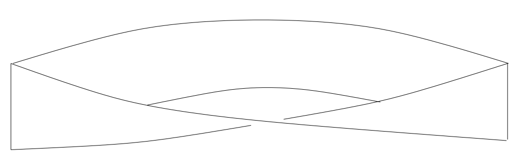
2.5 Normal vector fields and orientability

Definition 2.5.1
- *A surface is* orientable *if there is a (of course smooth) unit normal vector field* $\boldsymbol{U}$ *on* M.
- *A choice of unit normal vector field for a surface is called an* orientation *of the surface.*
- *An orientable surface with a prescribed unit normal vector field is called an* oriented *surface.*

NOTE: All the surfaces we will study in this course are orientable.

Remark 2.5.1

There are surfaces that are not orientable. The typical example is the Möbius strip.



Remark 2.5.2

Given a local patch σ on an oriented surface M, we have

$$\text{either} \quad \boldsymbol{U}_\sigma(u,v) = \frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|} \qquad \text{or} \quad \boldsymbol{U}_\sigma(u,v) = -\frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|}.$$

Example 2.5.1

For the surface parametrised by $\sigma(u,v) = (u, v, u^2 + v^2)$, $\sigma_u = (1, 0, 2u)$, $\sigma_v = (0, 1, 2v)$, so $\sigma_u \times \sigma_v = (-2u, -2v, 1)$ and therefore

$$\boldsymbol{U}_{\sigma(u,v)} = \frac{(-2u, -2v, 1)}{\sqrt{4u^2 + 4v^2 + 1}} \quad \text{or} \quad \boldsymbol{U}_{\sigma(u,v)} = \frac{(2u, 2v, -1)}{\sqrt{4u^2 + 4v^2 + 1}}.$$

Remark 2.5.3

For the rest of the course many things will depend on the choice of the unit normal vector. Both choices of normal are perfectly admissible and many of the quantities we will define using $\boldsymbol{U}$ will not be invariant under a change of orientation, so before an orientation is chosen there will be a degree of ambiguity in our definitions.

2.6 Differentiation on surfaces

To be able to study the shape of surfaces we will have to differentiate functions and vector fields that are only defined on a surface, and not on an open set of $\mathbb{R}^3$. In this and previous courses you have seen the definition of derivative for functions defined on an open set in $\mathbb{R}^n$, but not for functions defined in surfaces. Hence we have to define the derivative on surfaces.

The derivative measures the rate of change of some quantity when another quantity changes. Note that in a surface we have many directions in which to measure these rates of change, so we need to define derivatives with respect to a direction (actually with respect to a vector). These are called *directional derivatives.*

Definition 2.6.1 *Let M be a regular surface, $f : M \to \mathbb{R}^n$ a function and $X : M \to \mathbb{R}^3$ a vector field.*

The directional derivative *of f or X with respect to $\vec{w} \in T_pM$ is defined, respectively, as*

$$D_{\vec{w}} f(p) = \frac{d}{dt}\Big|_{t_0} f(\alpha(t)) \quad \text{or} \quad D_{\vec{w}} X_p = \frac{d}{dt}\Big|_{t_0} X(\alpha(t)),$$

where $\alpha : (a, b) \to M$ is a curve such that $\alpha(t_0) = p$ and $\alpha'(t_0) = \vec{w}$.

(Note that the definition for f and X is exactly the same; a distinction is made just because of the 'vector field character' of X: X_p is thought of as 'anchored' at p.)

How can we calculate these objects in a painless way?

Proposition 2.6.1 *Let $f : M \subset \mathbb{R}^3 \to \mathbb{R}^n$ be a function and X be a vector field in M.*

- *If f is defined not only on M but on an open subset of $\mathbb{R}^3$ containing M, then*

  $$D_{\vec{w}} f(p) = Df_p(\vec{w}),$$

  where Df_p is the usual derivative of f at p.

- *If $\vec{w} = a(\sigma_u)_p + b(\sigma_v)_p \in T_pM$ then*

$$D_{\vec{w}} X_p = a \left. \frac{\partial (X \circ \sigma)}{\partial u} \right|_{(u_0,v_0)} + b \left. \frac{\partial (X \circ \sigma)}{\partial v} \right|_{(u_0,v_0)},$$

where (u_0, v_0) are such that $p = \sigma(u_0, v_0)$. Note that in particular,

$$D_{\sigma_u} X_p = \left. \frac{\partial (X \circ \sigma)}{\partial u} \right|_{(u_0,v_0)} \quad \text{and} \quad D_{\sigma_v} X_p = \left. \frac{\partial (X \circ \sigma)}{\partial u} \right|_{(u_0,v_0)}.$$

PROOF :

- Let $\alpha : (a, b) \to M$ be a smooth curve with $\alpha(t_0) = p$, $\alpha'(t_0) = \vec{w}$. Using the chain rule,

$$D_{\vec{w}} f(p) = \left. \frac{d}{dt} \right|_{t_0} f(\alpha(t)) = Df_{\alpha(t_0)}(\alpha'(t_0)) = Df_p(\vec{w}).$$

- Consider the curve $\alpha(t) = \sigma(u_0 + at, v_0 + bt)$. Then $\alpha(0) = \sigma(u_0, v_0) = p$ and $\alpha'(0) = a(\sigma_u)_p + b(\sigma_v)_p = \vec{w}$. Hence, using the chain rule,

$$\begin{aligned} D_{\vec{w}} X_p &= \left. \frac{dX \circ \sigma(u_0 + at, v_0 + bt)}{dt} \right|_{t=0} \\ &= \left. \frac{\partial (X \circ \sigma)}{\partial u} \right|_{\alpha(0)} \frac{d(u_0 + at)}{dt} + \left. \frac{\partial (X \circ \sigma)}{\partial u} \right|_{\alpha(0)} \frac{d(v_0 + bt)}{dt} \\ &= a \left. \frac{\partial (X \circ \sigma)}{\partial u} \right|_{(u_0,v_0)} + b \left. \frac{\partial (X \circ \sigma)}{\partial v} \right|_{(u_0,v_0)}. \end{aligned}$$

$\square$

Some properties of the directional derivative

PROPOSITION 2.6.2 *The directional derivative satisfies*

- *For $\vec{w}_1, \vec{w}_2 \in T_pM$,*

$$D_{a_1\vec{w}_1 + a_2\vec{w}_2} X_p = a_1 D_{\vec{w}_1} X_p + a_2 D_{\vec{w}_2} X_p.$$

- *For X and Y vector fields,*

$$D_{\vec{w}}(X \cdot Y)(p) = (D_{\vec{w}} X_p) \cdot Y_p + X_p \cdot (D_{\vec{w}} Y_p).$$

This also holds if we substitute $\cdot$ with $\times$.

- *If σ is a surface patch about p, then*

$$(D_{\sigma_u} \sigma_v)_p = (D_{\sigma_v} \sigma_u)_p.$$

PROOF :

- Write $\vec{w}_1$ and $\vec{w}_2$ in the basis given by σ_u and σ_v, collect terms and apply the last statement of the previous proposition.

- If $\alpha : (a, b) \to M$ is such that $\alpha(t_0) = p$ and $\alpha'(t_0) = \vec{w}$ then

$$\begin{aligned} D_{\vec{w}}(X \cdot Y)(p) &= \left. \frac{X(\alpha(t)) \cdot Y(\alpha(t))}{dt} \right|_{t_0} = \left(\left. \frac{X(\alpha(t))}{dt} \right|_{t_0} \right) \cdot Y + X \cdot \left(\left. \frac{Y(\alpha(t))}{dt} \right|_{t_0} \right) \\ &= (D_{\vec{w}} X_p) \cdot Y_p + X_p \cdot (D_{\vec{w}} Y_p). \end{aligned}$$

The proof for $\times$ is the same.

- Use the last statement of the previous proposition:

$$(D_{\sigma_u}\sigma_v)_p = \frac{\partial(\sigma_v)_{\sigma(u,v)}}{\partial u} = \sigma_{uv} = \sigma_{vu} = \frac{\partial(\sigma_u)_{\sigma(u,v)}}{\partial v} = (D_{\sigma_v}\sigma_u)_p.$$

□

Chapter 3

Local geometry of surfaces

Now we have the tools from calculus we need in order to start studying the shape of surfaces.

For curves we differentiated the tangent vector in order to obtain the curvature. We could do the same in surfaces, except that we do not quite have a tangent vector but a tangent plane. Now, each tangent plane is completely characterised by the unit normal vector, so our strategy will be to differentiate the normal in order to obtain information about the shape of a surface.

3.1 The shape operator

Definition 3.1.1 *Let M be an regular oriented surface with unit normal $\boldsymbol{U}$. The* shape operator *on M is the linear transformation*

$$S_p : T_pM \to T_pM$$

defined by

$$S_p(\vec{w}) = -D_{\vec{w}}\boldsymbol{U}_p.$$

Remark 3.1.1

Note that S is actually something like a 'field' of linear transformations: S gives a linear map S_p for each $p \in M$.

Observe also that there is no reason a priori that $-D_{\vec{w}}\boldsymbol{U}_p$ will actually lie in T_pM. We prove this in the following lemma.

Lemma 3.1.1 *Let M be an oriented surface and $p \in M$. Then we have*

- *S_p is well defined as a map into T_pM in the sense that $S_p(\vec{w}) \in T_pM$.*
- *S_p is linear.*
- *S_p is symmetric, in the sense that $S_p(\vec{v}) \cdot \vec{w} = \vec{v} \cdot S_p(\vec{w})$.*

Proof :

- Since $\boldsymbol{U} \cdot \boldsymbol{U} = 1$, we have $0 = D_{\vec{w}}(\boldsymbol{U} \cdot \boldsymbol{U})(p) = 2(D_{\vec{w}}\boldsymbol{U}_p) \cdot \boldsymbol{U}_p$. Therefore $S_pD_{\vec{w}}\boldsymbol{U}_p$ is perpendicular to $\boldsymbol{U}_p$ so it lies in T_pM.
- Using the properties of the directional derivative, we have $S_p(a\vec{v} + b\vec{w}) = -D_{a\vec{v}+b\vec{w}}\boldsymbol{U} = --aD_{\vec{v}}\boldsymbol{U} - bD_{\vec{w}}\boldsymbol{U} = aS_p(\vec{v}) + bS_p(\vec{w})$.

- Let σ be a surface patch about p. Then we can write $\vec{v} = a_1(\sigma_u)_p + a_2(\sigma_v)_p$ and $\vec{w} = b_1(\sigma_u)_p + b_2(\sigma_v)_p$. Thus we have (we remove the subscript $_p$ in σ_u and σ_v for easier reading)

$$\begin{aligned} S_p(\vec{v}) \cdot \vec{w} &= S_p(a_1\sigma_u + a_2\sigma_v) \cdot (b_1\sigma_u + b_2\sigma_v) \\ &= (a_1 S_p(\sigma_u) + a_2 S_p(\sigma_v)) \cdot (b_1\sigma_u + b_2\sigma_v) \\ &= a_1b_1(S_p(\sigma_u)\cdot\sigma_u) + a_1b_2(S_p(\sigma_u)\cdot\sigma_v) + a_2b_1(S_p(\sigma_v)\cdot\sigma_u) \\ &\qquad + a_2b_2(S_p(\sigma_v)\cdot\sigma_v). \end{aligned}$$

Similarly,

$$\begin{aligned} \vec{v} \cdot S_p(\vec{w}) &= a_1b_1(S_p(\sigma_u)\cdot\sigma_u) + a_1b_2(S_p(\sigma_v)\cdot\sigma_u) + a_2b_1(S_p(\sigma_u)\cdot\sigma_v) \\ &\qquad + a_2b_2(S_p(\sigma_v)\cdot\sigma_v). \end{aligned}$$

Comparing the two expressions it becomes clear that it suffices to show that $S_p(\sigma_v)\cdot\sigma_u = S_p(\sigma_u)\cdot\sigma_v$. Now, observe that since $(\sigma_u \cdot \boldsymbol{U}) = 0$, we have

$$0 = D_{\sigma_u}(\sigma_v \cdot \boldsymbol{U})(p) = D_{\sigma_u}\sigma_v \cdot \boldsymbol{U}_p + \sigma_v \cdot D_{\sigma_u}\boldsymbol{U}_p = \sigma_{uv} \cdot \boldsymbol{U} + \sigma_v \cdot D_{\sigma_u}\boldsymbol{U}_p.$$

Similarly,

$$0 = D_{\sigma_v}(\sigma_u \cdot \boldsymbol{U})(p) = \sigma_{vu} \cdot \boldsymbol{U} + \sigma_u \cdot D_{\sigma_v}\boldsymbol{U}_p.$$

Since $\sigma_{uv} = \sigma_{vu}$ we finally have

$$S_p(\sigma_u)\cdot\sigma_v = -\sigma_v \cdot D_{\sigma_u}\boldsymbol{U} = \sigma_{uv}\cdot\boldsymbol{U} = -\sigma_u \cdot D_{\sigma_v}\boldsymbol{U}_p = S_p(\sigma_v)\cdot\sigma_u.$$

□

NOTATION: Given a surface patch σ, the matrix expression of S_p in the basis $\{(\sigma_u)_p, (\sigma_v)_p\}$ will be denoted by $[S_p]_\sigma$.

Definition 3.1.2 *The symmetric bilinear function*

$$\mathrm{II}_p : T_pM \times T_pM \to \mathbb{R}$$

defined by

$$\mathrm{II}_p(\vec{v}, \vec{w}) = S_p(\vec{v}) \cdot \vec{w}$$

is called the second fundamental form.

It is clear by the last statement of the last proposition that II_p is symmetric in the sense that $\mathrm{II}(\vec{v}, \vec{w}) = \mathrm{II}(\vec{w}, \vec{v})$.

Lemma 3.1.2 *Let σ be a surface patch about p. Then*

$$\mathrm{II}_p(a_1\sigma_u + a_2\sigma_v, b_1\sigma_u + b_2\sigma_v) = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} e & f \\ f & g \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

where

$$e = \sigma_{uu} \cdot \boldsymbol{U}_p \qquad f = \sigma_{uv} \cdot \boldsymbol{U}_p \qquad g = \sigma_{vv} \cdot \boldsymbol{U}_p.$$

Proof :

By bilinearity we have

$$\mathrm{II}_p(a_1\sigma_u + a_2\sigma_v, b_1\sigma_u + b_2\sigma_v) = a_1b_1\,\mathrm{II}(\sigma_u, \sigma_u) + (a_1b_2 + a_2b_1)\,\mathrm{II}(\sigma_u, \sigma_v) + a_2b_2\,\mathrm{II}(\sigma_v, \sigma_v),$$

and from the proof of Lemma 3.1.1 we know that

$$\mathrm{II}(\sigma_u, \sigma_u) = \sigma_{uu} \cdot \boldsymbol{U}_p \qquad \mathrm{II}(\sigma_v, \sigma_u) = \sigma_{uv} \cdot \boldsymbol{U}_p \qquad \mathrm{II}(\sigma_v, \sigma_v) = \sigma_{vv} \cdot \boldsymbol{U}_p.$$

so the only thing that remains to be checked is that

$$a_1 b_1\, e + (a_1 b_2 + a_2 b_1)\, f + a_2 b_2\, g = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} e & f \\ f & g \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

but this is straightforward.

□

NOTATION: For a surface patch σ, we will use the notation

$$[\mathrm{II}_p]_\sigma = \begin{pmatrix} e & f \\ f & g \end{pmatrix}.$$

(Note that the coefficients e, f, g certainly depend on the surface patch chosen.)

What is the relation between $[S_p]_\sigma$ and the matrix $[\mathrm{II}_p]_\sigma$? We need to make another very important definition that will also be very important when we talk about length and area in a surface.

Definition 3.1.3 *The symmetric bilinear function*

$$\mathrm{I}_p : T_pM \times T_pM \to \mathbb{R}$$

defined by

$$\mathrm{I}_p(\vec{v}, \vec{w}) = \vec{v} \cdot \vec{w}$$

is called the first fundamental form*.*

Lemma 3.1.3 *Let σ be a surface patch about p. Then*

$$\mathrm{I}_p(a_1\sigma_u + a_2\sigma_v, b_1\sigma_u + b_2\sigma_v) = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

where

$$E = \sigma_u \cdot \sigma_u \qquad F = \sigma_u \cdot \sigma_v \qquad G = \sigma_v \cdot \sigma_v.$$

Proof :

By bilinearity we have

$$\mathrm{I}_p(a_1\sigma_u + a_2\sigma_v, b_1\sigma_u + b_2\sigma_v) = a_1 b_1\, (\sigma_u \cdot \sigma_u) + (a_1 b_2 + a_2 b_1)\, (\sigma_u \cdot \sigma_v) + a_2 b_2\, (\sigma_v \cdot \sigma_v),$$

so the result follows.

□

Remark 3.1.2

It may look strange to give a new name and a new symbol to our good old dot product. The reason for this is that the dot product is defined in all of $\mathbb{R}^3$, whereas the first fundamental form is defined only on the tangent planes of the surface. So in a way is the gadget we would use if we lived in the surface without knowing we actually lived in $\mathbb{R}^3$. We will see more of this later.

NOTATION: For a surface patch σ, we will use the notation

$$[\mathrm{I}_p]_\sigma = \begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

(Note that the coefficients E, F, G certainly depend on the surface patch chosen.)

Proposition 3.1.1 *For σ a surface patch about p we have*

$$[\mathrm{II}_p]_\sigma = [\mathrm{I}_p]_\sigma \; [S_p]_\sigma.$$

Proof :

Let $\vec{v} = a_1\sigma_u + a_2\sigma_v$ and $\vec{w} = b_1\sigma_u + b_2\sigma_v$. Then, by definition,

$$\mathrm{II}_p(\vec{w}, \vec{v}) = \vec{w} \cdot S_p(\vec{v}) = \mathrm{I}_p(\vec{w}, S_p(\vec{v})).$$

In terms of matrices this can be written as

$$\begin{pmatrix} w_1 & w_2 \end{pmatrix} [\mathrm{II}_p]_\sigma \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} w_1 & w_2 \end{pmatrix} [\mathrm{I}_p]_\sigma \; [S_p]_\sigma \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Since the last equality holds for all w_1, w_2, v_1, v_2, the result follows.

□

Remark 3.1.3

The previous proposition is very useful for calculations: finding $[\mathrm{I}_p]_\sigma$ and $[\mathrm{II}_p]_\sigma$ is a routine matter, whereas finding $[S_p]_\sigma$ directly can be messier. So we will often use the formula

$$[S_p]_\sigma = [\mathrm{I}_p]_\sigma^{-1} \; [\mathrm{II}_p]_\sigma.$$

3.2 Curvature

It is maybe intuitively clear that the shape operator gives us a lot of information about the shape of a surface, as it says how the normal moves as we move along the surface. But S_p is a linear operator, and we would like to have some numbers that summarise the information that S_p may give. Certainly the matrix expression for S_p in a patch σ does not do the job since S_p does not depend on σ but $[S_p]_\sigma$ does, as the matrix of a linear transformation depends on the basis chosen. However, neither the determinant or the trace depend on the basis chosen, and furthermore they determine S_p completely. The eigenvalues and eigenvectors of S_p are also independent of any basis chosen. This motivates the following definitions.

Definition 3.2.1 *Let M be an oriented regular surface with shape operator S.*

- *The Gaussian curvature of M at p is defined by*

$$K(p) = \det S_p.$$

- *The mean curvature of M at p is defined by*

$$H(p) = \frac{\mathrm{Trace}(S_p)}{2}.$$

- *The principal curvatures at p are the eigenvalues of S_p.*
- *The principal directions at p are the eigenvectors of S_p.*
- *The asymptotic directions are the vectors $\vec{v} \in T_pM$ such that $\mathrm{II}_p(\vec{v}, \vec{v}) = 0$.*

Remark 3.2.1

- The linear transformation S_p can be diagonalised because it is symmetric. Furthermore, there is always an orthonormal basis of eigenvectors. This is a standard result in linear algebra.
- We therefore have that, if the principal curvatures are different, then **the principal directions are always perpendicular**.
- If the principal curvatures at p are equal then every vector in T_pM will be a principal direction. So in particular we could choose two perpendicular principal directions.

Before going into calculating all these quantities in a concrete example, let us derive some properties and formulas for the objects we just defined.

The following is a standard result in linear algebra.

Lemma 3.2.1 *Let $\vec{w} \in T_pM$ such that $\vec{w} \cdot \vec{w} = 1$ (or $\mathrm{I}_p(\vec{w}, \vec{w}) = 1$—it is the same thing). Let k_1 and k_2 denote the principal curvatures at p and assume that $k_2 \leq k_1$. Then*

$$k_2 \leq \mathrm{II}(\vec{w}, \vec{w}) \leq k_1.$$

Proof :

Let $\{\vec{v_1}, \vec{v_2}\}$ be an orthonormal basis of eigenvectors of S_p so that $\vec{v_1}$ has eigenvalue k_1 and $\vec{v_2}$ has eigenvalue k_2. We can then write $\vec{w} = a\vec{v_1} + b\vec{v_2}$ for some coefficients a, b satisfying $a^2 + b^2 = 1$ (since $\vec{w} \cdot \vec{w} = 1$). Then, by bilinearity and using the fact that $S_p(\vec{v_1}) = k_1\vec{v_1}$ and $S_p(\vec{v_2}) = k_2\vec{v_2}$, we have

$$\begin{aligned} \mathrm{II}_p(\vec{w}, \vec{w}) &= a^2\, S_p(\vec{v}_1) \cdot \vec{v}_1 + 2ab\, S_p(\vec{v}_1) \cdot \vec{v}_2 + b^2\, S_p(\vec{v}_2) \cdot \vec{v}_2 \\ &= a^2 k_1 \vec{v}_1 \cdot \vec{v}_1 + 2abk_1\vec{v}_1 \cdot \vec{v}_2 + b^2 k_2 \vec{v}_2 \cdot \vec{v}_2 \\ &= k_1 a^2 + k_2 b^2. \end{aligned}$$

Now, since $k_2 \leq k_1$,

$$k_2 = k_2(a^2 + b^2) = k_2 a^2 + k_2 b^2 \leq k_1 a^2 + k_2 b^2 \leq k_1 a^2 + k_1 b^2 = k_1(a^2 + b^2) = k_1,$$

as desired.

□

Proposition 3.2.1 *(Some formulas)*

- *If k_1 and k_2 are the principal curvatures then $K(p) = k_1 k_2$ and $H(p) = (k_1 + k_2)/2$.*
- *If K and H are the Gaussian and mean curvatures, respectively, then the principal curvatures are given by $H \pm \sqrt{H^2 - K}$.*
- *If σ is a surface patch and $[\mathrm{I}_p]_\sigma = \begin{pmatrix} E & F \\ F & G \end{pmatrix}$ and $[\mathrm{II}_p]_\sigma = \begin{pmatrix} e & f \\ f & g \end{pmatrix}$ then*

$$K(p) = \frac{eg - f^2}{EG - F^2} \qquad \textit{and} \qquad H(p) = \frac{\mathrm{Trace}(S_p)}{2} = \frac{Ge + Eg - 2Ff}{2(EG - F^2)}.$$

Proof :

- In a basis given by principal directions (i.e. eigenvectors of S_p), the matrix of S_p has the form $\begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}$, so the result follows.

- Solve the system of equations given by $K = k_1k_2$ and $H = (k_1 + k_2)/2$: multiply H by $2k_1$ to obtain $2Hk_1 = k_1^2 + k_1k_2 = k_1^2 + K$. Thus $k_1^2 - 2Hk_1 + K = 0$, so completing the square we get $(k_1 - H)^2 + K = H^2$, and the result follows.

- Since $[S_p]_\sigma = [\mathrm{I}_p]_\sigma^{-1} [\mathrm{II}]_\sigma$, we have $K(p) = \det([S_p]_\sigma) = \det([\mathrm{I}_p]_\sigma^{-1}) \det([\mathrm{II}]_\sigma) = (eg - f^2)/(EF - G^2)$.

  To find the formula for H, we have

$$\begin{aligned}
[S_p]_\sigma &= [\mathrm{I}_p]_\sigma^{-1} [\mathrm{II}_p]_\sigma \\
&= \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} e & f \\ f & g \end{pmatrix} \\
&= \frac{1}{EG - F^2} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix} \begin{pmatrix} e & f \\ f & g \end{pmatrix} \\
&= \frac{1}{EG - F^2} \begin{pmatrix} Ge - Ff & Gf - Fg \\ Ef - Fe & Eg - Ff \end{pmatrix}.
\end{aligned}$$

  Therefore

$$H(p) = \frac{\mathrm{Trace}(S_p)}{2} = \frac{Ge + Eg - 2Ff}{2(EG - F^2)}.$$

□

Example 3.2.1

Let us calculate all these objects for the torus of external radius R and internal radius r, with $R > r$.

Recall that a patch for the torus that covers most of it is

$$\sigma(u, v) = ((R + r\cos v)\cos u, (R + r\cos v)\sin u, r\sin v), \quad 0 < u, v < 2\pi.$$

We have

$$\sigma_u = (R + r\cos v)(-\sin u, \cos u, 0) \qquad \sigma_v = (-r\sin v\cos u, -r\sin v\sin u, r\cos v)$$

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

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

Therefore a choice of normal unit vector is

$$\boldsymbol{U} = \frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|} = \frac{r(R + r\cos v)(\cos u\cos v, \sin u\cos v, \sin v)}{r(R + r\cos v)} = (\cos u\cos v, \sin u\cos v, \sin v).$$

Denoting by p a generic point $\sigma(u, v)$ we have

$$S_p(\sigma_u) = -D_{\sigma_u}\boldsymbol{U} = -\frac{\partial \boldsymbol{U}}{\partial u} = (\sin u\cos v, -\cos u\cos v, 0) = \frac{-\cos v}{R + r\cos v}\sigma_u$$

$$S_p(\sigma_v) = -D_{\sigma_v}\boldsymbol{U} = -\frac{\partial \boldsymbol{U}}{\partial v} = (\sin v\cos u, \sin v\sin u, -\cos v) = \frac{-\sigma_v}{r}.$$

Therefore the matrix expression of S_p in the basis $\{\sigma_u, \sigma_v\}$ is

$$[S_p]_\sigma = \begin{pmatrix} \dfrac{-\cos v}{R + r\cos v} & 0 \\ 0 & \dfrac{-1}{r} \end{pmatrix}.$$

Therefore we have

$$K(p) = \frac{\cos v}{r(R + r\cos v)} \qquad H(p) = \frac{-\cos v}{2(R + r\cos v)} + \frac{-1}{2r}.$$

Since $[S_p]_\sigma$ is already diagonalised, the principal curvatures are

$$\frac{-\cos v}{2(R + r\cos v)} \quad \text{with principal direction } \sigma_u \text{ and } \quad \frac{-1}{2r} \quad \text{with principal direction } \sigma_v.$$

Now we calculate $[\mathrm{I}_p]_\sigma$. We have $E = \sigma_u \cdot \sigma_u = (R+r\cos v)^2$, $F = \sigma_u \cdot \sigma_v = 0$ and $G = \sigma_v \cdot \sigma_v = r^2$. Therefore

$$[\mathrm{I}_p]_\sigma = \begin{pmatrix} (R + r\cos v)^2 & 0 \\ 0 & r^2 \end{pmatrix}.$$

And now we find $[\mathrm{II}_p]_\sigma$. We have $e = \sigma_{uu} \cdot \boldsymbol{U} = -(R + r\cos v)\cos v$, $f = \sigma_{uv} \cdot \boldsymbol{U} = 0$ and $g = \sigma_{vv} \cdot \boldsymbol{U} = -r$. Therefore

$$[\mathrm{II}_p]_\sigma = \begin{pmatrix} -(R + r\cos v)\cos v & 0 \\ 0 & -r \end{pmatrix}.$$

Note that the formula $[\mathrm{I}_p]_\sigma\ [S_p]_\sigma = [\mathrm{II}]_\sigma$ is clearly satisfied.

Finally we find the asymptotic directions. Let $\vec{w} = a\sigma_u + b\sigma_v$. Then

$$\mathrm{II}(\vec{w}, \vec{w}) = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} -(R + r\cos v)\cos v & 0 \\ 0 & -r \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = -a^2(R + r\cos v)\cos v - b^2 r.$$

Therefore we have to solve the equation $(R + r\cos v)\cos v a^2 + rb^2 = 0$. This has solutions only when $(R + r\cos v)\cos v < 0$, i.e. when $\cos v \le 0$, which means that $\pi \le v \le 3\pi/2$. In this case, $b = \pm a\sqrt{-(R + r\cos v)\cos v/r}$. Therefore the asymptotic directions are

$$\sigma_u \pm \sqrt{-(R + r\cos v)\cos v/r}\ \sigma_v.$$

So far we have only done the calculations without further interpretation of the quantities we obtain. We will do this in due course, but first we need to introduce some more concepts.

3.3 Length and area

Suppose that you are a 2-dimensional being living in a 2-dimensional surface M in $\mathbb{R}^3$. First of all, there is no apparent way that you could know that you live in a space of higher than 2 dimensions. All your study of your universe would be based in measurements made in the surface where you live. How much will you be able to know about your world?

Since we are (only!) 3-dimensional beings, we can hopefully help these 'poor' creatures. They can certainly measure angles and lengths on the surface, so they do know the first fundamental form I at each point p of their world. In fact this is why it is called 'first fundamental form': it is exactly what you can measure if you are a 2D living in the surface. However they do not know that this very fundamental form actually comes from a dot product in a space of 3 dimensions.

It is clear that these 2D beings can measure lengths of curves: if α is a curve in the surface, you just have to integrate $\sqrt{\mathrm{I}(\alpha'(t), \alpha'(t))}$.

Now, can they measure area? Yes: break the surface in small rectangles, measure the lengths of the sides of each rectangle, multiply the lengths and add up over all the rectangles. What is an analytic formula for this?

Recall the following definition.

Definition 3.3.1 *Let M be a surface and let $\sigma : U \to M$ be a surface patch. Then*

$$\mathit{Area}(\sigma(U)) = \iint_U \|\sigma_u \times \sigma_v\|\, dv\, du.$$

Remark 3.3.1 If M is covered by a single surface patch σ except maybe for some small subsets of M (such as curves in the surface) then $\text{Area}(M) = \text{Area}(\sigma(U))$, so the previous definition applies.

The interesting fact is that area is independent of parametrization, and in fact we can find a formula in terms of the first fundamental form only:

Proposition 3.3.1 *Let $\sigma : U \to M$ be a surface patch. Then*

$$\|\sigma_u \times \sigma_v\| = \sqrt{\det([\mathrm{I}]_\sigma)}.$$

Therefore, using E, F and G to denote the entries of the matrix $[\mathrm{I}]_\sigma$ as usual,

$$\mathit{Area}(\sigma(U)) = \iint_U \sqrt{EG - F^2}\, dv\, du.$$

Proof :

Recall the formula $(\vec{u} \times \vec{v}) \times \vec{w} = (\vec{u} \cdot \vec{w})\, \vec{v} - (\vec{v} \cdot \vec{w})\, \vec{u}$ and the formula $(\vec{u} \times \vec{v}) \cdot \vec{w} = (\vec{w} \times \vec{u}) \cdot \vec{v}$. We have

$$\begin{aligned}(\sigma_u \times \sigma_v) \cdot (\sigma_u \times \sigma_v) &= ((\sigma_u \times \sigma_v) \times \sigma_u) \cdot \sigma_v = ((\sigma_u \cdot \sigma_u)\, \sigma_v - (\sigma_u \cdot \sigma_v)\, \sigma_u) \cdot \sigma_v \\ &= (\sigma_u \cdot \sigma_u)(\sigma_v \cdot \sigma_v) - (\sigma_u \cdot \sigma_v)(\sigma_u \cdot \sigma_v) = EG - F^2.\end{aligned}$$

□

3.4 Curves in surfaces

We continue with our discussion of two dimensional beings in M. They can certainly measure velocity. They can also measure acceleration, but they only feel those components of acceleration that lie in their 2 dimensional world. In other words, 2D's do not feel any components of acceleration that are normal to the surface, they do not exist for them. We know they do exist, but let us see what the 2D's would measure as acceleration.

Definition 3.4.1 *Let $\alpha(t)$ be a regular curve in an oriented regular surface M with normal $\boldsymbol{U}$. Let $\boldsymbol{T}(t)$ and $\boldsymbol{N}(t)$ be the tangent and normal vectors of α and let $\kappa(t)$ be its curvature. The function*

$$\kappa_n(t) = \kappa(t)\, \boldsymbol{N}(t) \cdot \boldsymbol{U}_{\alpha(t)}$$

is called the normal curvature *of the curve α in M, and the function*

$$\kappa_g(t) = \kappa(t)\, \boldsymbol{N}(t) \cdot (\boldsymbol{U}_{\alpha(t)} \times \boldsymbol{T}(t))$$

is called the geodesic curvature *of the curve α in M.*

Proposition 3.4.1

$$\kappa(t)\boldsymbol{N}(t) = \kappa_n(t)\, \boldsymbol{U}_{\alpha(t)} + \kappa_g(t)\, (\boldsymbol{U}_{\alpha(t)} \times \boldsymbol{T}(t)).$$

Proof : Let $p = \alpha(t)$ and denote $\boldsymbol{T}(t)$ simply by $\boldsymbol{T}$, etc. Note that the vectors $\boldsymbol{U}$, $\boldsymbol{T}$ and $\boldsymbol{U} \times \boldsymbol{T}$ form an orthonormal basis ($\boldsymbol{U}$ and $\boldsymbol{T}$ are perpendicular because $\boldsymbol{T} \in T_pM$ and $\boldsymbol{U}$ is perpendicular to T_pM). Note that $\boldsymbol{U} \times \boldsymbol{T} \in T_pM$ since it is perpendicular to $\boldsymbol{U}$.

Since $\boldsymbol{N} \cdot \boldsymbol{T} = 0$, we have

$$\kappa\boldsymbol{N} = (\kappa\boldsymbol{N} \cdot \boldsymbol{T})\, \boldsymbol{T} + (\kappa\boldsymbol{N} \cdot \boldsymbol{U})\, \boldsymbol{U} + (\kappa\boldsymbol{N} \cdot (\boldsymbol{U} \times \boldsymbol{T}))\, \boldsymbol{U} \times \boldsymbol{T} = \kappa_n\boldsymbol{U} + \kappa_g(\boldsymbol{U} \times \boldsymbol{T}).$$

□

Corollary 3.4.1

$$\kappa^2 = \kappa_n^2 + \kappa_g^2.$$

Proof :

Clear since $\{\boldsymbol{U}, \boldsymbol{T}, \boldsymbol{U} \times \boldsymbol{T}\}$ form an orthonormal basis.

□

Remark 3.4.1

Suppose that a 2D inhabitant of M rides her motorcycle at constant speed 1 along some curve in M. What centrifugal forces will he feel?

We can parametrise this movement by a curve α parametrised by arc length. We then have

$$\alpha''(s) = \kappa(s)\boldsymbol{N}(s) = \kappa_n(s)\,\boldsymbol{U}_{\alpha(s)} + \kappa_g(s)\,(\boldsymbol{U}_{\alpha(s)} \times \boldsymbol{T}(s)).$$

However, 2D inhabitants of M **do not feel normal directions** because the normal is in an extra dimension they do not comprehend. Therefore they will never feel the normal component of the centripetal acceleration $\kappa_n\boldsymbol{U}$; they will only feel the tangent component $\kappa_g(\boldsymbol{U} \times \boldsymbol{T})$.

In other words:

- κ_n is the normal component of acceleration, **not felt** by 2D's living in the surface.
- κ_g is the tangent component of acceleration, **felt** by 2D's living in the surface.

In particular, a curve in M will be the equivalent of a straight line if it has no acceleration at all as measured from the point of view of the surface, i.e. if $\kappa_g = 0$.

In order to relate these concepts with the principal curvatures we need the following very important fact.

Proposition 3.4.2 *Let α be a regular curve in an oriented surface M with normal $\boldsymbol{U}$. Then the normal curvature of α in M is given by*

$$\kappa_n(t) = \mathrm{II}_{\alpha(t)}(\boldsymbol{T}(t), \boldsymbol{T}(t)),$$

where $\boldsymbol{T}(t)$ is the tangent vector to α.

Proof :

Since $\alpha'(t) \in T_{\alpha(t)}M$, $\alpha'(t) \cdot \boldsymbol{U}_{\alpha(t)} = 0$. Differentiating and applying the definition of directional derivative for $\boldsymbol{U}$ we obtain

$$\alpha''(t) \cdot \boldsymbol{U}_{\alpha(t)} + \alpha'(t) \cdot D_{\alpha'(t)}\boldsymbol{U}_{\alpha(t)} = 0.$$

Note that by definition $-\alpha'(t) \cdot D_{\alpha'(t)}\boldsymbol{U}_{\alpha(t)} = \alpha'(t) \cdot S_{\alpha(t)}(\alpha'(t)) = \mathrm{II}_{\alpha(t)}(\alpha'(t), \alpha'(t))$.

Recall that we have

$$\alpha''(t) = Sp'(t)\boldsymbol{T}(t) + [Sp(t)]^2\kappa(t)\boldsymbol{N}(t).$$

Since $\boldsymbol{T} \cdot \boldsymbol{U} = 0$ we have

$$\alpha''(t) \cdot \boldsymbol{U}_{\alpha(t)} = [Sp(t)]^2\,\kappa(t)\;\boldsymbol{N}(t) \cdot \boldsymbol{U}_{\alpha(t)},$$

and therefore, since $Sp(t) = \|\alpha'(t)\|$ and II_p is bilinear,

$$\kappa_n(t) = \frac{\mathrm{II}_{\alpha(t)}(\alpha'(t), \alpha'(t))}{[Sp(t)]^2} = \mathrm{II}_{\alpha(t)}\left(\frac{\alpha(t)}{\|\alpha(t)\|}, \frac{\alpha(t)}{\|\alpha(t)\|}\right) = \mathrm{II}_{\alpha(t)}(\boldsymbol{T}(t), \boldsymbol{T}(t)).$$

□

Proposition 3.4.3 *If α is a regular curve in an oriented surface M and $k_1 \geq k_2$ are the principal curvatures of M then at every point of the curve we have*

$$k_2 \leq \kappa_n \leq k_1.$$

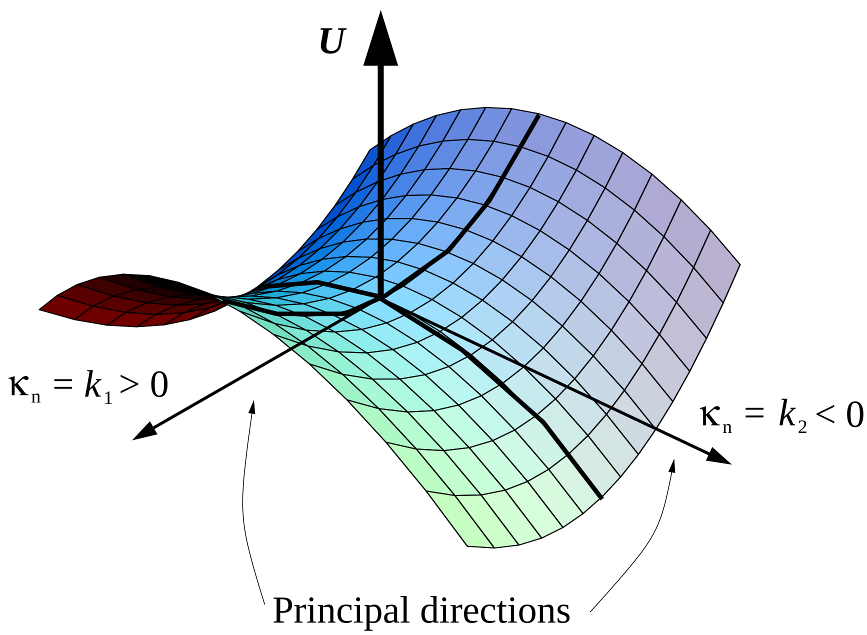
Proof :

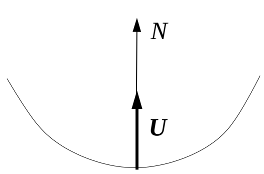
Immediate from Lemma 3.2.1 and Proposition 3.4.2.

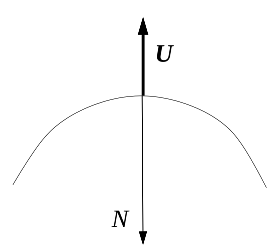
□

Remark 3.4.2

The last proposition tells us what the principal curvatures are: they are the (normal) curvatures of the most curved curves in the surface, with positive or negative sign depending of how they curve with respect to the normal:







$\kappa_n > 0$ (N and $\boldsymbol{U}$ point in the same direction)

$\kappa_n < 0$ (N and $\boldsymbol{U}$ point in the opposite directions)

Definition 3.4.2 *Let M be a regular oriented surface. A regular curve $\alpha : (a,b) \to M$ is*

- *A* line of curvature *if $\alpha'(t)$ is a principal direction at every point $\alpha(t)$.*

- *An* asymptotic line *if* $\mathrm{II}_{\alpha(t)}(\alpha'(t), \alpha'(t)) = 0$ *for all* t.
- *A* geodesic *if* $\alpha''(t)$ *is zero or parallel to the normal vector* $\boldsymbol{U}$ *at* $\alpha(t)$ *for all* t.

Remark 3.4.3

- Lines of curvature are therefore those curves in which the surface curves the most either upwards with respect to the normal ($k_n > 0$) or downwards with respect to the normal ($k_n < 0$).
- Geodesics are curves whose acceleration points in the direction of the normal to the surface. From the point of view of the 2 dimensional inhabitants of M these curves have no acceleration at all. We see this in the following proposition.

Proposition 3.4.4 *A regular curve* α *is a geodesic if and only if it has constant speed and its geodesic curvature* $\kappa_g = 0$.

Proof :

If α is a geodesic then $\alpha''(t)$ is parallel to $\boldsymbol{U}_{\alpha(t)}$. Therefore it is perpendicular to T_pM and in particular it is perpendicular to $\alpha'(t) \in T_pM$. This implies

$$\frac{d(\alpha'(t) \cdot \alpha'(t))}{dt} = 2(\alpha'(t) \cdot \alpha''(t)) = 0.$$

Therefore $(\alpha'(t) \cdot \alpha'(t))$ is constant, so α has constant speed, which we will denote by Sp.

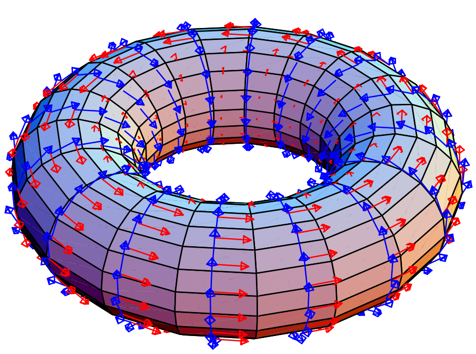
This implies in particular that $\alpha''(t) = Sp^2\, \kappa(t)\, \boldsymbol{N}(t)$. Hence $\boldsymbol{N}(t)$ is parallel to $\alpha''(t)$ and therefore to $\boldsymbol{U}_{\alpha(t)}$, and we have

$$\kappa_g(t) = \kappa(t)\, \boldsymbol{N}(t) \cdot (\boldsymbol{U}_{\alpha(t)} \times \boldsymbol{T}(t)) = 0.$$

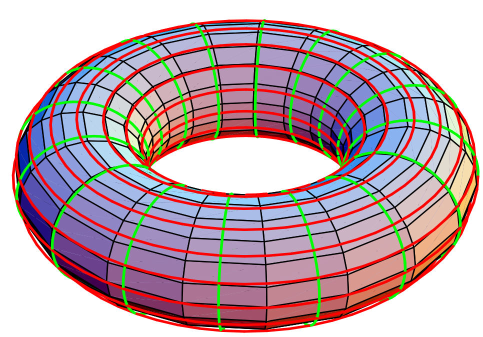
Now we prove the converse. Since $\kappa_g(t) = \kappa(t)\, \boldsymbol{N}(t) \cdot (\boldsymbol{U}_{\alpha(t)} \times \boldsymbol{T}(t)) = 0$ and $\boldsymbol{N}$ is perpendicular to $\boldsymbol{T}$, we must have that $\boldsymbol{N}(t)$ is parallel to $\boldsymbol{U}_{\alpha(t)}$. On the other hand, α has constant speed Sp, so $\alpha''(t) = Sp\, \kappa(t)\, \boldsymbol{N}(t)$. Thus, since $\boldsymbol{N}(t)$ is parallel to $\boldsymbol{U}_{\alpha(t)}$, $\alpha''(t)$ must also be parallel to $\boldsymbol{U}_{\alpha(t)}$.

$\square$

The following pictures show the principal directions and the lines of curvature for the torus.

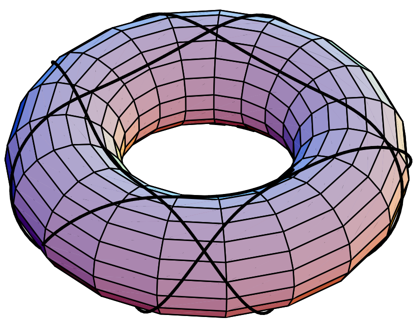


Principal directions on the torus

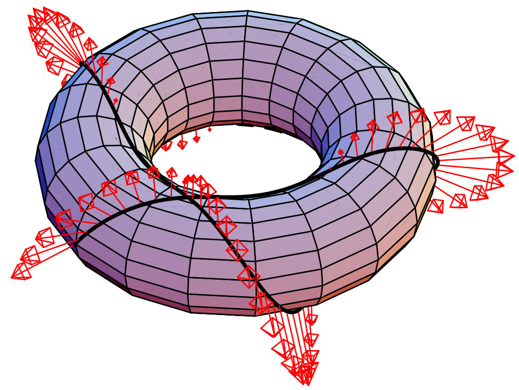


Lines of curvature on the torus

The picture on the left is a geodesic on the torus. The one on the right has the plot of a geodesic together with its normal $\boldsymbol{N}$ (actually $-\boldsymbol{N}$ is depicted so that it points to the outside of the surface). Note that $\boldsymbol{N}$ is always perpendicular to the surface, as it should be.



Geodesic in torus



Geodesic and normal $-\boldsymbol{N}$ to curve

3.5 Local equations for geodesics

How do we find the geodesics in a surface?

Proposition 3.5.1 *Let M be a regular oriented surface with normal $\boldsymbol{U}$ and let $\sigma : U \subset \mathbb{R}^2 \to M$ be a surface patch. Then the curve $\alpha(t) = \sigma(\beta_1(t), \beta_2(t))$ is a geodesic if and only if it satisfies the system of differential equations*

$$\begin{aligned} 2E\ \beta_1'' + E_u\ [\beta_1']^2 + 2E_v\ \beta_1'\beta_2' + (2F_v - G_u)\ [\beta_2']^2 + 2F\ \beta_2'' &= 0 \\ 2F\ \beta_1'' + (2F_u - E_v)\ [\beta_1']^2 + 2G_u\ \beta_1'\beta_2' + G_v\ [\beta_2']^2 + 2G\ \beta_2'' &= 0 \end{aligned}$$

where E, F and G are the coefficients of the first fundamental form in the patch σ.

Proof :

The curve $\alpha(t)$ is a geodesic if and only if $\alpha''(t)$ is zero or parallel to $\boldsymbol{U}_{\alpha(t)}$; equivalently $\alpha(t)$ is a geodesic if and only if $\alpha''(t)$ is zero or perpendicular to the coordinate vectors σ_u and σ_v, i.e. if and only if $\alpha''(t) \cdot \sigma_u = \alpha''(t) \cdot \sigma_v = 0$ at the point $\alpha(t)$ for all t.

Now,

$$\alpha'(t) = \beta_1'(t)\, \sigma_u(\beta_1(t), \beta_2(t)) + \beta_2'(t)\, \sigma_v(\beta_1(t), \beta_2(t)),$$

and

$$\begin{aligned} \alpha''(t) \quad = \quad & \beta_1''(t)\, \sigma_u(\beta_1(t), \beta_2(t)) + [\beta_1'(t)]^2\, \sigma_{uu}(\beta_1(t), \beta_2(t)) + \beta_1'(t)\beta_2'(t)\, \sigma_{uv}(\beta_1(t), \beta_2(t)) \\ & + \beta_2''(t)\, \sigma_v(\beta_1(t), \beta_2(t)) + [\beta_2'(t)]^2\, \sigma_{vv}(\beta_1(t), \beta_2(t)) + \beta_1'(t)\beta_2'(t)\, \sigma_{uv}(\beta_1(t), \beta_2(t)). \end{aligned}$$

Using the relations

$$\sigma_u \cdot \sigma_u = E, \quad \sigma_{uu} \cdot \sigma_u = E_u/2, \quad \sigma_{uv} \cdot \sigma_u = E_v/2, \quad \sigma_v \cdot \sigma_u = F, \quad \sigma_{vv} \cdot \sigma_u = F_v - G_u/2$$

$$\sigma_{uu} \cdot \sigma_v = F_u - E_v/2, \quad \sigma_{vv} \cdot \sigma_v = G_v/2, \quad \sigma_{uv} \cdot \sigma_v = G_u/2, \quad \sigma_v \cdot \sigma_v = G$$

we find

$$2\alpha''(t) \cdot \sigma_u(\beta_1(t), \beta_2(t)) = 2E\beta_1'' + E_u[\beta_1']^2 + 2E_v\beta_1'\beta_2' + (2F_v - G_u)[\beta_2']^2 + 2F\beta_2''$$

and

$$2\alpha''(t) \cdot \sigma_v(\beta_1(t), \beta_2(t)) = 2F\beta_1'' + (2F_u - E_v)[\beta_1']^2 + 2G_u\beta_1'\beta_2' + G_v[\beta_2']^2 + 2G\beta_2''$$

Hence $\alpha = \sigma(\beta_1(t), \beta_2(t))$ is a geodesic if and only if β_1 and β_2 satisfy the following system of second order differential equations.

$$\begin{aligned} 2E\ \beta_1'' + E_u\ [\beta_1']^2 + 2E_v\ \beta_1'\beta_2' + (2F_v - G_u)\ [\beta_2']^2 + 2F\ \beta_2'' &= 0 \\ 2F\ \beta_1'' + (2F_u - E_v)\ [\beta_1']^2 + 2G_u\ \beta_1'\beta_2' + G_v\ [\beta_2']^2 + 2G\ \beta_2'' &= 0 \end{aligned}$$

□

In general it is difficult to solve these equations. Note that, being a system of 2 second order equations, we need 4 initial conditions, normally the value of the curve α at some point and its velocity at that point.

Facts about geodesics: We list some important and interesting facts that we shall not prove.

- Given two different points in a surface there is always a geodesic passing through these points.

- Among all the geodesics through two given points, there is one that has shortest length. In fact, this shortest geodesic **gives the shortest path to get from one point to the other** while staying in the surface. In other words, it is for the 2D inhabitants of the surface what a straight line is for us.

3.6 An interpretation of Gaussian curvature

Gaussian curvature turns out to be the main invariant of surfaces, in the sense that two surfaces with the same first fundamental form will have the same Gaussian curvature (we will see this fact later). But we have not quite given a geometric interpretation of curvature yet.

Note that, although the shape operator, the second fundamental form, the principal curvatures and the mean curvature change sign when we change the orientation of the surface, Gaussian curvature does not since the determinant of a 2 by 2 matrix is equal to the determinant of minus that matrix.

Definition 3.6.1 *Let M be a regular surface. A point p in M is called*

- *Elliptic if $K(p) > 0$.*

- *Hiperbolic if $K(p) < 0$.*

- *Parabolic if $K(p) = 0$ but one of the principal curvatures at p is not 0 (or equivalently S_p is not 0).*

- *Planar if the principal curvatures at p are both 0 (or equivalently S_p is 0).*

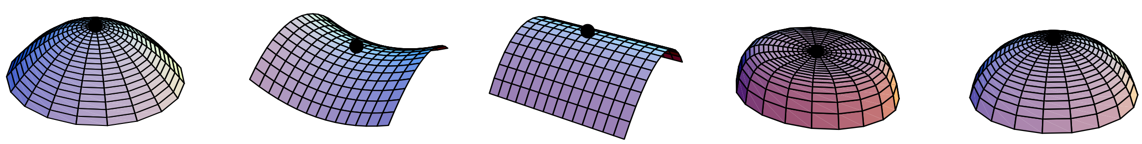
- *Umbillic if the principal curvatures at p are equal.*

Facts:

From our interpretation of the principal directions as the directions in T_pM in which the surface curves the most and the principal curvatures as the signed magnitude of this curving, positive if it curves towards the normal and negative if it curves away from the normal, we have the following facts. Of course all the statements that follow are local, i.e. in a small neighbourhood of the point in question.

- At an elliptic point all the curves in the surface through that point curve in the same direction, either towards the normal to the surface or away from it.
- At a hyperbolic point some curves in the surface through that point curve towards the normal to the surface and others away from it.
- At a parabolic point all curves in the surface through that point curve towards the normal to M or away from it except for one that does not curve at all with respect to the normal to M.
- At a planar point curves in the surface through that point do not curve at all with respect to the normal to M.
- At an umbillic point all curves in the surface through that point curve the same amount in the same direction with respect to the normal to M.

In pictures:



Elliptic Hyperbolic Parabolic Planar Umbillic

The following theorem is similar in flavour to some theorems we saw about curves.

Theorem 3.6.1 *Let M be a regular surface. If all the points of M are umbillic, then M is either contained in a sphere or in a plane.*

Proof :

We will only prove that the theorem is true if it is covered by a single surface patch σ; to prove it in general requires some knowledge about topology.

Since p is umbillic, the eigenvalues of S_p at every point p are equal; therefore every vector in T_pM must be an eigenvector, so both σ_u and σ_v are eigenvectors with eigenvalue $k_1(=k_2)$. Thus

$$\frac{\partial \boldsymbol{U}(\sigma(u,v))}{\partial u} = k_1(\sigma(u,v))\,\frac{\partial \sigma(u,v)}{\partial u} \qquad \text{and} \qquad \frac{\partial \boldsymbol{U}(\sigma(u,v))}{\partial v} = k_1(\sigma(u,v))\,\frac{\partial \sigma(u,v)}{\partial v}.$$

Differentiating the first equation with respect to v and the second with respect to u we obtain

$$\frac{\partial^2 \boldsymbol{U}(\sigma(u,v))}{\partial v \partial u} = \frac{\partial k_1(\sigma(u,v))}{\partial v}\frac{\partial \sigma(u,v)}{\partial u} + k_1(\sigma(u,v))\,\frac{\partial^2 \sigma(u,v)}{\partial v \partial u}$$

and

$$\frac{\partial^2 \boldsymbol{U}(\sigma(u,v))}{\partial u \partial v} = \frac{\partial k_1(\sigma(u,v))}{\partial u}\frac{\partial \sigma(u,v)}{\partial u} + k_1(\sigma(u,v))\,\frac{\partial^2 \sigma(u,v)}{\partial u \partial v}.$$

Since second derivatives commute we have

$$\frac{\partial k_1(\sigma(u,v))}{\partial v}\,\sigma_u = \frac{\partial k_1(\sigma(u,v))}{\partial u}\,\sigma_v,$$

and since σ_u and σ_v are linearly independent, we must have

$$\frac{\partial k_1(\sigma(u,v))}{\partial v} = \frac{\partial k_1(\sigma(u,v))}{\partial u} = 0,$$

proving that k_1 is constant.

Now we integrate the equations

$$\frac{\partial \boldsymbol{U}(\sigma(u,v))}{\partial u} = k_1\,\frac{\partial \sigma(u,v)}{\partial u} \qquad \text{and} \qquad \frac{\partial \boldsymbol{U}(\sigma(u,v))}{\partial v} = k_1\,\frac{\partial \sigma(u,v)}{\partial v},$$

obtaining

$$\boldsymbol{U}(\sigma(u,v)) = k_1\sigma(u,v) + \vec{p}_0$$

for some constant vector $\vec{p}_0$.

If $k_1 = 0$ then $\boldsymbol{U}$ is constant. Let $\sigma(u_0,v_0)$ be a point in M, and consider the function

$$g(u,v) = (\sigma(u,v) - \sigma(u_0,v_0))\cdot \boldsymbol{U}.$$

Then, since $\boldsymbol{U}$ is constant, $g_u(u,v) = \sigma_u \cdot \boldsymbol{U} = 0$ and $g_v(u,v) = \sigma_v \cdot \boldsymbol{U} = 0$, which implies that $f(u,v)$ is constant. Since $f(u_0,v_0) = 0$ we must have

$$f(u,v) = (\sigma(u,v) - \sigma(u_0,v_0))\cdot \boldsymbol{U} = 0 \quad \text{for all } u, v,$$

which expresses the fact that the image of σ lies in the plane through $\sigma(u_0,v_0)$ perpendicular to the constant vector $\boldsymbol{U}$.

If $k_1 \neq 0$, then

$$\sigma(u,v) + \frac{\vec{p}_0}{k_1} = \frac{\boldsymbol{U}(\sigma(u,v))}{k_1},$$

and therefore

$$\left\|\sigma(u,v) - \left(-\frac{\vec{p}_0}{k_1}\right)\right\| = \frac{1}{|k_1|}$$

which means that the image of σ lies in a sphere centred at $-\vec{p}_0/k_1$ with radius $1/k_1$.

□

3.7 Example: Surfaces of revolution

A surface of revolution is the surface obtained when a curve is rotated around an axes. Suppose that the rotation axes is the z-axes and let the curve be $(0, r(s), h(s))$ for $s \in (a,b)$; without loss of generality we assume that the curve is parametrised by arc length (so $(r')^2 + (h')^2 = 1$). For the surface to be regular we need to assume that $r > 0$.

A patch of the surface is

$$\sigma(u,v) = (r(v)\cos u, r(v)\sin u, h(v)), \quad 0 < u < 2\pi, a < v < b.$$

σ covers all the surface except for a curve, but changing the domain of σ all the surface gets covered. So we will do all calculations using σ.

We have

$$\sigma_u = (-r(v)\sin u, r(v)\cos, 0) \qquad\qquad \sigma_v = (r'(v)\cos u, r'(v)\sin u, h'(v)).$$

Therefore

$$[\mathrm{I}]_\sigma = \begin{pmatrix} [r(v)]^2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{so} \quad E = [r(v)]^2, F = 0, G = 1.$$

We have $\sigma_u \times \sigma_v = r(v)\ (h'(v)\cos u, h'(v)\sin u, -r'(v))$; this has norm $r(v)$, so we will take

$$\boldsymbol{U} = (h'(v)\cos u, h'(v)\sin u, -r'(v))$$

as the unit normal. The matrix expression of the second fundamental form (recall that the coefficients are given by $e = \sigma_{uu} \cdot \boldsymbol{U}$, $f = \sigma_{uu} \cdot \boldsymbol{U}$, $g = \sigma_{vv} \cdot \boldsymbol{U}$) is

$$[\mathrm{II}]_\sigma = \begin{pmatrix} -r(v)h'(v) & 0 \\ 0 & r''(v)h'(v) - r'(v)h''(v) \end{pmatrix}.$$

Since $(r')^2 + (h')^2 = 1$, $r'r'' = -h'h''$, so $r''h' - r'h'' = r''h' + (r')^2 r''/h' = r''/h'$. Thus,

$$[\mathrm{II}]_\sigma = \begin{pmatrix} -r(v)h'(v) & 0 \\ 0 & r''(v)/h'(v) \end{pmatrix}.$$

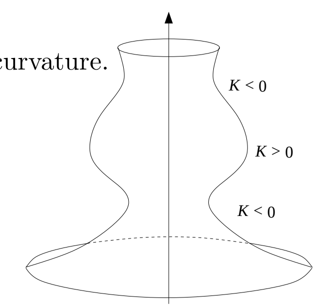
Therefore the matrix expression for the shape operator in the patch σ is

$$[S]_\sigma = [\mathrm{I}]_\sigma^{-1}\ [\mathrm{II}]_\sigma = \begin{pmatrix} -\dfrac{h'(v)}{r(v)} & 0 \\ 0 & r''(v)/h'(v) \end{pmatrix}.$$

We immediately obtain that

- The principal directions are σ_u and σ_v with corresponding principal curvatures $-h'(v)/r(v)$ and $r''(v)h'(v) - r'(v)h''(v)$.
- Therefore the parallels $\sigma(t, v_0)$ and the meridians $\sigma(u_0, t)$ are lines of curvature.
- The gaussian curvature is $K(\sigma(u,v)) = -\dfrac{r''(v)}{r(v)}$.

  This means that the sign of K depends only on the concavity or convexity of $r(v)$.



Now we find the geodesics. We have

$$E = [r(v)]^2, \quad F = 0, \quad G = 1, \quad E_u = 0, \quad E_v = 2r(v)r'(v), \quad F_u = F_v = G_u = G_v = 0.$$

We use Proposition 3.5.1. we have that $\alpha(t) = \sigma(\beta_1(t), \beta_2(t))$ is a geodesic if and only if (to simplify the notation, let us write $\beta_1(t) = u$ and $\beta_2(t) = v$; we will not write the argument 't' but we should not forget that u and v are functions of t, and that all the dashes are derivatives with respect to t; the subscripts $_v$ mean derivatives with respect to v.)

$$\begin{aligned} 2[r(v)]^2\ u'' + 4\, r(v)\ r_v(v)\ u'\ v' &= 0 \\ -r(v)\ r_v(v)\ [u']^2 + 2v'' &= 0 \end{aligned}$$

The first equation can be written as

$$(\,[r(v)]^2\ u'\,)' = 0,$$

and therefore

$$[r(v)]^2\ u' = c$$

for some constant c. Now we know that a geodesic has constant speed Sp, so

$$\alpha'(t) \cdot \alpha'(t) = (u'\sigma_u + v'\sigma_v) \cdot (u'\sigma_u + v'\sigma_v) = E(u')^2 + 2Fu'v' + G(v')^2 = Sp^2.$$

Therefore,

$$[r(v)]^2(u')^2 + (v')^2 = Sp^2.$$

Now substitute $[r(v)]^2\, u' = c$ in the previous equation to obtain

$$c^2 + [r(v)]^2(v')^2 = Sp^2[r(v)]^2.$$

This equation can be integrated to obtain v and then the equation $[r(v)]^2\, u' = c$ can be integrated to obtain u.

The main interesting point in this discussion is the following. It is clear from the parametrization that σ_u is tangent to a parallel of the surface (i.e. σ_u is parallel to the xy-plane). Now observe that

$$\alpha'(t) \cdot \sigma_u = (u'\sigma_u + v'\sigma_v) \cdot \sigma_u = Eu' = [r(v)]^2\, u' = c.$$

Since $\alpha'(t) \cdot \sigma_u = \|\alpha'(t)\|\, \|\sigma_u\| \cos\theta = Sp\ r(v) \cos\theta$, where θ is the angle between $\alpha'(t)$ and a parallel, we obtain $Sp\ r(v) \cos\theta = c$. Since Sp is constant, we can rewrite this as (**Clairaut's relation**):

$$r(v) \cos\theta = \text{constant}.$$

In other words, **the cosine of the angle that a geodesic makes with the horizontal is inversely proportional to the distance to the axis of rotation**.

There are some geodesics that are easy to find. The **meridians** $\sigma(u_0, t)$ **are geodesics** since in this case we have $u' = 0$ and $v' = 1$, so the system of equations above is clearly satisfied.

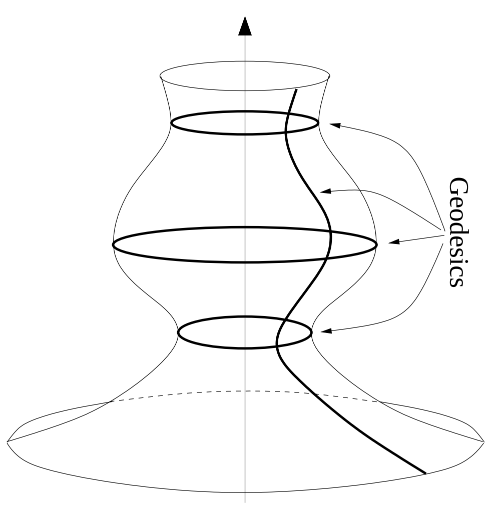
Are the parallels $\alpha(t) = \sigma(t, v_0)$ geodesics? Not in general. Let us see the conditions for a parallel to be a geodesic. Plugging in $u = t$, $v = v_0$ in the equations above we see that the first equation is satisfied (since both u'' and v' are 0). Since $u' = 1$ and $v'' = 0$, the second equation becomes

$$-r(v_0)\, r'(v_0) = 0.$$

Since r is never 0, this implies that a parallel $\sigma(t, v_0)$ is a geodesic if and only if $r'(v_0) = 0$; in other words, the only parallels of a surface of revolution that are geodesics are the 'waists' and the 'hips'.

It is immediate to check that any meridian parametrised with constant speed is also a geodesic as the equations are clearly satisfied (plug in, for example, $u = u_0$ and $v = t$).

In pictures:



3.8 Gauss *Theorema Egregium*

In this section we look at Gauss' remarkable theorem that says that the Gaussian curvature only depends on the first fundamental form. This is extraordinary: it means that by means of measurements only on the surface one should be able to figure out how curved the surface is. In other words, the 2 dimensional inhabitants of a surface are able to find out, for example, whether they are standing in an elliptic, hyperbolic o parabolic point.

Recall that for curves we differentiated the tangent in order to obtain information about the shape of the curve. So far for surfaces we have differentiated the normal. One reason for this is that there is not one tangent vector but a whole plane of them.

Let us try to do a similar analysis. Take a coordinate patch σ in a regular oriented surface M. Then $\{\sigma_u, \sigma_v, \boldsymbol{U}\}$ form a basis of $\mathbb{R}^3$ at each point $\sigma(u,v)$ of the surface. This is kind of the Frenet frame that we had for surfaces except for the very important fact that it is not an orthonormal basis.

Let us differentiate these three vectors in the direction of σ_u and σ_v as we did in the case of curves for the Frenet frame. First, the normal is easy:

$$D_{\sigma_u}\boldsymbol{U} = -s_{11}\sigma_u - s_{21}\sigma_v, \quad D_{\sigma_v}\boldsymbol{U} = -s_{12}\sigma_u - s_{22}\sigma_v,$$

where the s_{ij} are the coefficients of the matrix representation of the shape operator in the basis σ_u and σ_v, i.e.

$$[S]_\sigma = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}.$$

The other derivatives can be written as

$$\begin{aligned} D_{\sigma_u}\sigma_u &= \Gamma^1_{11}\sigma_u + \Gamma^2_{11}\sigma_v + L_1\boldsymbol{U} \\ D_{\sigma_u}\sigma_v &= \Gamma^1_{12}\sigma_u + \Gamma^2_{12}\sigma_v + L_2\boldsymbol{U} \\ D_{\sigma_v}\sigma_v &= \Gamma^1_{22}\sigma_u + \Gamma^2_{22}\sigma_v + L_3\boldsymbol{U} \end{aligned}$$

Where $\Gamma^1_{11}, \Gamma^2_{11}, \Gamma^1_{12}, \Gamma^2_{12}, \Gamma^1_{22}, \Gamma^2_{22}$ are some coefficients, called the *Christoffel symbols*. We do not calculate $D_{\sigma_v}\sigma_u$ because it is equal to $D_{\sigma_u}\sigma_v$.

Note that $L_1 = (D_{\sigma_u}\sigma_u)\cdot\boldsymbol{U} = -\sigma_u \cdot D_{\sigma_u}\boldsymbol{U} = \mathrm{II}(\sigma_u,\sigma_u) = e$ in the traditional notation. Similarly, $L_2 = f$ and $L_3 = g$. So our complete set of equations is

$$\begin{aligned} D_{\sigma_u}\sigma_u &= \Gamma^1_{11}\sigma_u + \Gamma^2_{11}\sigma_v + e\boldsymbol{U} \\ D_{\sigma_u}\sigma_v &= \Gamma^1_{12}\sigma_u + \Gamma^2_{12}\sigma_v + f\boldsymbol{U} \\ D_{\sigma_v}\sigma_v &= \Gamma^1_{22}\sigma_u + \Gamma^2_{22}\sigma_v + g\boldsymbol{U} \\ D_{\sigma_u}\boldsymbol{U} &= -s_{11}\sigma_u - s_{21}\sigma_v \\ D_{\sigma_v}\boldsymbol{U} &= -s_{12}\sigma_u - s_{22}\sigma_v \end{aligned}$$

To find the Christoffel symbols, dot-multiply each equation by σ_u and σ_v to obtain

$$\begin{aligned} \Gamma^1_{11}E + \Gamma^2_{11}F &= \sigma_{uu}\cdot\sigma_u = E_u/2 \\ \Gamma^1_{11}F + \Gamma^2_{11}G &= \sigma_{uu}\cdot\sigma_v = F_u - E_v/2 \\ \Gamma^1_{12}E + \Gamma^2_{12}F &= \sigma_{uv}\cdot\sigma_u = E_v/2 \\ \Gamma^1_{12}F + \Gamma^2_{12}G &= \sigma_{uv}\cdot\sigma_v = G_u/2 \\ \Gamma^1_{22}E + \Gamma^2_{22}F &= \sigma_{vv}\cdot\sigma_u = F_v - G_u/2 \\ \Gamma^1_{22}F + \Gamma^2_{22}G &= \sigma_{vv}\cdot\sigma_v = G_v/2 \end{aligned}$$

where we have used the relations

$$\sigma_u \cdot \sigma_u = E, \quad \sigma_{uu} \cdot \sigma_u = E_u/2, \quad \sigma_{uv} \cdot \sigma_u = E_v/2, \quad \sigma_v \cdot \sigma_u = F, \quad \sigma_{vv} \cdot \sigma_u = F_v - G_u/2$$

$$\sigma_{uu} \cdot \sigma_v = F_u - E_v/2, \quad \sigma_{vv} \cdot \sigma_v = G_v/2, \quad \sigma_{uv} \cdot \sigma_v = G_u/2, \quad \sigma_v \cdot \sigma_v = G$$

The point is that we have six equations for six variables $\Gamma^1_{11}, \Gamma^2_{11}, \Gamma^1_{12}, \Gamma^2_{12}, \Gamma^1_{22}, \Gamma^2_{22}$. Therefore **the Christoffel symbols can be obtained in terms of the coefficients** E, F, G **of the first fundamental form**. Now it turns out that the curvature can be found in terms of these symbols, as we will now see.

We complicate things further: note that since multiple derivatives commute we have relations of the sort

$$\begin{aligned} D_{\sigma_v} D_{\sigma_u} \sigma_u &= \sigma_{vuu} = \sigma_{uvu} = D_{\sigma_u} D_{\sigma_v} \sigma_u \\ D_{\sigma_v} D_{\sigma_v} \sigma_u &= \sigma_{vvu} = \sigma_{vuv} = D_{\sigma_v} D_{\sigma_u} \sigma_v \\ D_{\sigma_v} D_{\sigma_u} \boldsymbol{U} &= D_{\sigma_u} D_{\sigma_v} \boldsymbol{U} \end{aligned}$$

and many more (although it turns out that the rest of the relations do not add any more information—these are it!). These equations can be rewritten as

$$\begin{aligned} 0 &= D_{\sigma_v} D_{\sigma_u} \sigma_u - D_{\sigma_u} D_{\sigma_v} \sigma_u \\ 0 &= D_{\sigma_v} D_{\sigma_v} \sigma_u - D_{\sigma_v} D_{\sigma_u} \sigma_v \\ 0 &= D_{\sigma_v} D_{\sigma_u} \boldsymbol{U} - D_{\sigma_u} D_{\sigma_v} \boldsymbol{U}. \end{aligned}$$

If now we plug in the original equations into these ones (formidable task!) we will get many relations. It turns out that all the possible relations get summarised into three equations. The first is the following, called the *Gauss equation*. The subscripts u or v mean, as usual, derivatives with respect to u or v:

$$(\Gamma^2_{12})_u - (\Gamma^2_{11})_v + \Gamma^1_{12}\Gamma^2_{11} + \Gamma^2_{12}\Gamma^2_{12} - \Gamma^2_{11}\Gamma^2_{22} - \Gamma^1_{11}\Gamma^2_{12} = -E\frac{eg - f^2}{EG - F^2} = -E\,K.$$

This implies in particular that K **can be obtained solely from the coefficients of the first fundamental form and no further information**. This is

Theorem 3.8.1 (Gauss Theorema Egregium) *The Gaussian curvature K of a surface depends only on the first fundamental form. Therefore if we change the surface while preserving lengths and angles in the surface the curvature does not change.*

The two other relations are the *Codazzi-Mainardi* equations:

$$\begin{aligned} e_v - f_u &= e\Gamma^1_{12} + f(\Gamma^2_{12} - \Gamma^1_{11} - g\Gamma^2_{11}) \\ f_v - g_u &= e\Gamma^1_{22} + f(\Gamma^2_{22} - \Gamma^1_{12} - g\Gamma^2_{12}) \end{aligned}$$

Remember that for curves we had a theorem that said that given functions $\kappa > 0$ and τ there was a unique curve with curvature κ and torsion τ up to rigid motion. For surfaces we have an equivalent theorem:

Theorem 3.8.2 (Bonnet) *Given functions E, F, G and e, f, g defined in an open set in $\mathbb{R}^2$ with $E > 0$, $G > 0$ and $EG - F^2 > 0$ that satisfy the Gauss formula and the Codazzi-Mainardi equations there is a parametrised surface $\sigma(V) \subset \mathbb{R}^3$ whose coefficients of the first and second fundamental form are the given E, F, G and e, f, g; furthermore, any two surfaces with these characteristics differ only by a rigid motion.*