Visual interactive differential geometry

Lisbeth Fajstrup and Martin Raussen

October 2, 2015

Contents

1 Curves in Plane and Space

1.1 Vector functions and parametrized curves

You have certainly encountered curves before. Simple examples are straight lines, circles, ellipses; more complex ones the trajectory of a spacecraft or of a charged particle in an electromagnetic field. We want to describe a framework that allows us to describe and investigate general curves in plane and space.

1.1.1 Parametrization by vector functions

Many plane curves can be described as the graph of a function $f : [a,b] \to \mathbf{R}$. But such a simple curve as a plane circle cannot! And for space curves, it is obvious that one has to find other means of description.

Illustration 1.1 *graph of a function, circle, and space curve - do it yourself...*

We let $\mathbf{R}^n = \{[x_1, x_2, \ldots, x_n] |\ x_i \in \mathbf{R}$ denote ordinary n-dimensional space equipped with the orthonormal basis $\{\mathbf{e}_i | 1 \leq i \leq n\}$, $\mathbf{e}_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$

Definition 1.2 *Let $I \subset \mathbf{R}$ denote an open interval.*

1. *Any function* $\mathbf{r} : I \to \mathbf{R}^n$ *is called a* vector function. *A vector function can be described in coordinates as*

$$\mathbf{r}(t) = \sum_i r_i(t)\mathbf{e}_i = [r_1(t), \ldots, r_n(t)],\ t \in I.$$

2. *A vector function* $\mathbf{r} : I \to \mathbf{R}^n,\ \mathbf{r}(t)[r_1(t), \ldots, r_n(t)]$ *is called* smooth *(C^∞) if the coordinate functions $r_i(t)$ are infinitely many times differentiable on I. Its derivative* $\mathbf{r}' : I \to \mathbf{R}^n$ *is defined as* $\mathbf{r}'(t) = [r_1'(t), \ldots, r_n'(t)],\ t \in I.$

Remark 1.3 *An interval J can be of the form $]a, b[$, $]a, b]$, $[a, b[$ or $[a, b]$ for real numbers a, b. One can extend Def. 1.2 and define curves on intervals, which are not open - by requiring* $\mathbf{r}$ *to be continuous on J and smooth on the interior. We may also consider intervals of the form $] - \infty, b[$, $] - \infty, b]$, $]a, \infty[$, $]a, \infty]$ and $] - \infty, \infty[= \mathbf{R}$. The interior $int(J)$ removes boundary points from J; f.ex., $int([a, b]) =]a, b[$. This is needed to extend Def. 1.2 to general intervals, since differentiability makes only sense on* open *intervals!*

1.1.2 Parametrized curves

Let us get started with some examples: A line in Euclidean n-space through P in direction $\mathbf{x} \in \mathbf{R}^n \setminus \{\mathbf{0}\}$ can be parameterized by the vector function

$$\mathbf{r} : \mathbf{R} \to \mathbf{R}^n, \quad \mathbf{r}(t) = \overrightarrow{OP_t} = \overrightarrow{OP} + t\mathbf{x}, \quad t \in \mathbf{R}.$$

The vector function $\mathbf{c} : [0, 2\pi] \to \mathbf{R}^2$, $\mathbf{c}(t) = [\cos t, \sin t]$, represents a *circle* C in the Euclidean plane with radius 1 and the origin as its center: The circle consists of all points P_t with $\overrightarrow{OP_t} = \mathbf{c}(t)$. In both cases, you may imagine the arrow $\overrightarrow{OP_t}$ pointing at P_t *at time* t.

Remark, that a circle *cannot* be represented as the graph of a function $f : \mathbf{R} \to \mathbf{R}$: There are *two* elements $y = \pm\sqrt{1 - x^2}$ corresponding to an element $x \in (-1, 1)$ with $[x, y] \in C$. In space, it is even less reasonable to represent curves as graphs of functions.

Example 1.4 *The vector function* $\mathbf{r} : \mathbf{R} \to \mathbf{R}^3$, $\mathbf{r}(t) = [a\cos t, a\sin t, bt]$ *represents a helix winding around a cylinder of radius a with the z-axis as the central axis – above, resp. below a circle of radius a. The helix will be used as one of our central examples throughout this chapter.*

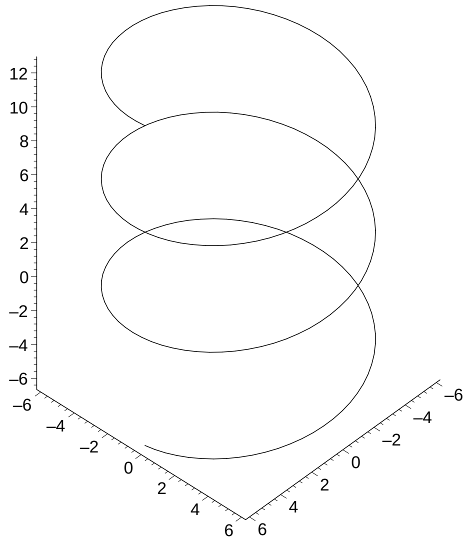


Figure 1.1: A helix

Illustration 1.5 *Better have an applet here: An interval, the circle/helix, and a scrollbar allowing to scroll a point one the interval, the corresponding point on the helix, and an arrow from the origin to that point (moving when scrolled!) See the geometric lab.*

Definition 1.6 *A smooth vector function* $\mathbf{r} : I \to \mathbf{R}^n$ *is called a* parametrization *of the curve*

$$C = \{P_t \in \mathbf{R}^n | \overrightarrow{OP_t} = \mathbf{r}(t),\, t \in I\}$$

The curve C consists thus of all the points P_t "pointed to" by arrows $\overrightarrow{OP_t} = \mathbf{r}(t)$. Often, you will imagine I as a time interval, and the curve given by a "particle" at position P_t at time t.

Example 1.7 (important general *example): Is it always possible to represent a the* graph of a function $f : I \to \mathbf{R}$ *in the plane by a (vector function) parametrization? Yes! Here is the* recipe*:*

The curve $C = \{[x, f(x)] \in \mathbf{R}^2 | x \in I\}$ *can be parameterized by the vector function* $\mathbf{r} : I \to \mathbf{R}^2$, $\mathbf{r}(t) = [t, f(t)]$, $t \in I$. *Then, the points* $P_t : [t, f(t)]$ *run through all the points on the curve* C*. For instance,* $\mathbf{r}(t) = [t, \sin t]$ *is a parametrization of the graph of the sine-function.*

Illustration 1.8 *Scrollable interval and arrows to the graph.*

1.1.3 Fundamental properties of smooth vector functions

Since we shall express geometric properties of curves using vector functions and their derivatives (up to degree 3), a brief investigation into properties of such vector functions will prove to be a good investment. The following rules apply to vector functions and ordinary functions that can be described as composites (sums, products, dot products, cross=wedge-products, composite of maps) of vector functions (and ordinary functions). Make sure that you understand the meaning of each term!

Proposition 1.9 *Let* $I, J \subseteq \mathbf{R}$ *denote intervals; moreover, let* $\mathbf{r}_1, \mathbf{r}_2 : I \to \mathbf{R}^n$ *denote smooth vector functions, and let* $f : I \to \mathbf{R}$ *and* $s : J \to I$ *denote (ordinary) smooth functions. The derivatives of compound functions satisfy the following rules at every* $t \in int(I)$*:*

1. $(\mathbf{r}_1 \pm \mathbf{r}_2)'(t) = \mathbf{r}_1'(t) \pm \mathbf{r}_2'(t)$;
2. $(f\mathbf{r}_1)'(t) = f'(t)\mathbf{r}_1(t) + f(t)\mathbf{r}_1'(t)$;
3. $(\mathbf{r}_1 \cdot \mathbf{r}_2)'(t) = \mathbf{r}_1'(t) \cdot \mathbf{r}_2(t) + \mathbf{r}_1(t) \cdot \mathbf{r}_2'(t)$;
4. $(\mathbf{r}_1 \times \mathbf{r}_2)'(t) = \mathbf{r}_1'(t) \times \mathbf{r}_2(t) + \mathbf{r}_1(t) \times \mathbf{r}_2'(t)$;
5. *(The chain rule)* $(\mathbf{r}_1 \circ s)'(t) = s'(t)\mathbf{r}_1'(s(t))$, $t \in J$.

Proof: Having described vector functions by their coordinates, the proofs are straightforward implications of the rules for derivatives of ordinary smooth functions. Note, that the rules for the dot product (3.) and the cross product (4.) have the form of the ordinary product rule, again. Note that the terms in (3.) are ordinary smooth functions!

□

1.1.4 An important consequence

The following consequence of the derivation rules above is completely elementary, but nevertheless a technically very important device:

Proposition 1.10 *(The fundamental trick).*
Let $c \in \mathbf{R}$ *denote a constant, and* $I \subseteq \mathbf{R}$ *an interval.*

1. *Let* $\mathbf{r}_1, \mathbf{r}_2 : I \to \mathbf{R}^n$ *denote two smooth vector functions with*

$$\mathbf{r}_1(t) \cdot \mathbf{r}_2(t) = c \textit{ for every } t \in I.$$

Then, $\mathbf{r}_1'(t) \cdot \mathbf{r}_2(t) + \mathbf{r}_1(t) \cdot \mathbf{r}_2'(t) = 0$ *for every* $t \in int(I)$.

2. *Let* $\mathbf{r} : I \to \mathbf{R}^n$ *denote a smooth vector function with constant length* $|\mathbf{r}(t)| = c$ *for every* $t \in I$. *Then* $\mathbf{r}(t) \cdot \mathbf{r}'(t) = 0$ *for every* $t \in int(I)$.
   In particular, $\mathbf{r}'(t)$ *is* perpendicular *on* $\mathbf{r}(t)$ *for every* $t \in int(I)$.

Proof: Apply the product rule (1.9.3) for vector functions.

□

1.2 Tangents, velocity, and length

1.2.1 Secants and tangents

For the graph of a function $f : I \to \mathbf{R}$, the *secant* through $[x_0, f(x_0)]$ and $[x_1, f(x_1)]$ with $x_0 \neq x_1$ is defined as the line through these two points on the graph. The *tangent* to the graph at $[x_0, f(x_0)]$ is defined as the limit position of these secant lines when x_1 tends to x_0 – if such a limit position exists. We know that it exists when f is smooth; in that case it is the line through $[x_0, f(x_0)]$ with slope $f'(x_0)$.

It is not very difficult to generalise these concepts to curves (in the plane, in 3-space or in n-space) given by a smooth parametrization. In fact, smoothness is not quite enough to guarantee the existence of a tangent line at every point. This can be seen in

Example 1.11 *Let C denote the curve given by the parametrization* $\mathbf{r} : \mathbf{R} \to \mathbf{R}^2$, $\mathbf{r}(t) = [1 + t^2, 1 + t^3]$, *cf. Fig. 1.2. Apparently, this curve has a singularity at the point* $\overrightarrow{OP_0} = [1, 1]$. *The secant lines do have a limit position, which is the horizontal line through* P_0. *But this line is not a good approximation to the curve near* P_0. *Points in the 2. quadrant in the line do not approximate the curve.*

Illustration 1.12 *Scroll on the tangent line and on the curve using the differential, both at* P_0 *and at* P_1.

But in most cases, a tangent line yields a good approximation, e.g., at the point $P_1 : [2, 2]$ *for the curve above. Why do we find different behaviour? Let us calculate derivatives:*

$$\mathbf{r}'(t) = [2t, 3t^2]; \quad \mathbf{r}'(0) = \mathbf{0} \quad \mathbf{r}'(1) = [1, 1] \neq \mathbf{0}!$$

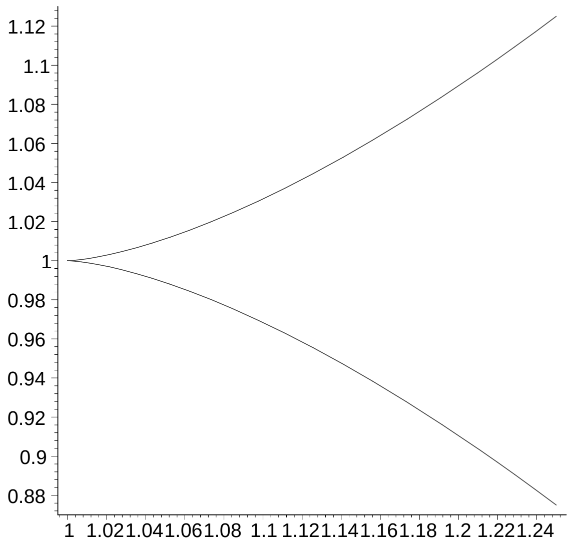


Figure 1.2: A curve with a singular point

Definition 1.13 1. *A smooth vector function* $\mathbf{r} : I \to \mathbf{R}^n$ *is called a* regular *parametrization of the curve* $C = \{P_t \in \mathbf{R}^n | \overrightarrow{OP_t} = \mathbf{r}(t), t \in I\}$ *if and only if* $\mathbf{r}'(t) \neq \mathbf{0}$ *for all* $t \in int(I)$.

2. *A subset* $C \subset \mathbf{R}^n$ *is called a* regular curve *if there is a regular parametrization* $\mathbf{r} : I \to \mathbf{R}^n$ *with* $C = \{P_t \in \mathbf{R}^n | \overrightarrow{OP_t} = \mathbf{r}(t), t \in I\}$.

Remark 1.14 1. *There is nothing like* the *parametrization of a curve. You can move along a curve at different speed, and speed need not be constant.*

2. *A regular curve may have a non-regular parametrization. This is the case for the circle with the parametrization* $\mathbf{c}(t) = [\cos t^2, \sin t^2], t \in \mathbf{R}$. *Why?*

3. *The curve C from Ex. 1.11 does not possess any regular parametrization.*

Tangents of regular curves

Given a (regular) curve C with a regular parametrization $\mathbf{r} : I \to \mathbf{R}^n$, it is easy to show that it has a tangent line at every point $P_t \in C$, $t \in int(I)$ and to determine it: Given a point $P_{t_0} \in C$ with $\mathbf{r}(t_0) = \overrightarrow{OP_{t_0}}$, $t_0 \in int(I)$. To get the tangent line to C at P_{t_0}, we only need a parallel vector as a limit vector from suitable parallel vectors of secant lines nearby. The secant between P_{t_0} and P_t with $\mathbf{r}(t) = \overrightarrow{OP_t}$ has parallel vector $\overrightarrow{P_{t_0}P_t} = \overrightarrow{OP_t} - \overrightarrow{OP_{t_0}} = \mathbf{r}(t) - \mathbf{r}(t_0)$ (for $P_{t_0} \neq P_t$). For $t \to t_0$, this difference vector gets shorter and shorter and thus tends to $\mathbf{0}$, hence does not give information on the direction of a possible tangent. Instead, we look at the *unit* parallel vector $\frac{\overrightarrow{P_{t_0}P_t}}{|\overrightarrow{P_{t_0}P_t}|} = \frac{\mathbf{r}(t)-\mathbf{r}(t_0)}{|\mathbf{r}(t)-\mathbf{r}(t_0)|}$ and its limit position for $t \to t_0$.

Illustration 1.15 *Applet with secants and tangents – See the geometric lab..*

Proposition 1.16 *Let* $\mathbf{r} : [a,b] \to \mathbf{R}^n$ *denote a* regular *parametrization of a curve* C. *Let* $t_0 \in (a,b)$ *and* $\overrightarrow{OP_{t_0}} = \mathbf{r}(t_0)$. *Then, the curve has a tangent line at* P_{t_0} *with parallel vector* $\mathbf{r}'(t_0)$. *A parametrization for the tangent line is* $\mathbf{t}(t) = \mathbf{r}(t_0) + t\mathbf{r}'(t_0)$.

Proof: For a regular parametrization, we take the limit of the unit direction vectors for $t \to t_0+$ from the left:

$$\lim_{t\to t_0+} \frac{\overrightarrow{P_{t_0}P_t}}{|\overrightarrow{P_{t_0}P_t}|} = \lim_{t\to t_0+} \frac{\mathbf{r}(t)-\mathbf{r}(t_0)}{|\mathbf{r}(t)-\mathbf{r}(t_0)|} = \frac{\lim_{t\to t_0+} \frac{\mathbf{r}(t)-\mathbf{r}(t_0)}{t-t_0}}{\lim_{t\to t_0+} \frac{|\mathbf{r}(t)-\mathbf{r}(t_0)|}{t-t_0}} = \frac{\mathbf{r}'(t_0)}{|\mathbf{r}'(t_0)|},$$

and similarly,

$$\lim_{t\to t_0-} \frac{\overrightarrow{P_{t_0}P_t}}{|\overrightarrow{P_{t_0}P_t}|} = -\frac{\mathbf{r}'(t_0)}{|\mathbf{r}'(t_0)|.}$$

(Why does this not work for $\mathbf{r}'(t_0) = \mathbf{0}$?). In particular, these two limit vectors are the two unit vectors parallel to $\mathbf{r}'(t_0)$. Hence the line through P_{t_0} with parallel vector $\mathbf{r}'(t_0)$ is a limit for the secant lines through P_{t_0} – and thus rightly qualifies as the tangent line.

□

Illustration 1.17 *Applet firing tangents from points at a regular curve given by a parametrization after choice. Including coordinates of position and direction.*

Remark that a regular curve has a tangent vector at each point. On the other hand, a curve with a non-regular parametrization may have tangent lines, but Ex. 1.11 shows that there may be points without a tangent line (regardless the parametrization). In many cases, one can still define semi-tangents using the limits

$$\lim_{t\to t_0+} \frac{\mathbf{r}(t)-\mathbf{r}(t_0)}{|\mathbf{r}(t)-\mathbf{r}(t_0)|}, \text{ resp. } \lim_{t\to t_0-} \frac{\mathbf{r}(t)-\mathbf{r}(t_0)}{|\mathbf{r}(t)-\mathbf{r}(t_0)|}$$

as the definition. See also Exc. 1.2.3.?.

1.2.2 Velocity and arc length

Speed and unit tangent vectors

Proposition 1.16 shows that the derivative $\mathbf{r}'(t)$ of a regular parametrization of a curve C contains the geometric information needed to determine the tangent lines to C. But there is more (non-geometric) information hidden in the vector function $\mathbf{r}'(t)$. If you imagine $\mathbf{r}(t)$ tracing the trajectory of a particle at time t, differentation should have something to do with the *speed* of that motion, as well.

How can one grasp the speed of a smooth vector function? As for functions of one variable, we look at the quotient $\frac{|\overrightarrow{P_{t_0}P_t}|}{|t-t_0|} = \frac{|\mathbf{r}(t)-\mathbf{r}(t_0)|}{|t-t_0|}$. For $t \to t_0$, we obtain the limit $\lim_{t\to t_0} \frac{|\mathbf{r}(t)-\mathbf{r}(t_0)|}{|t-t_0|} = |\lim_{t\to t_0} \frac{\mathbf{r}(t)-\mathbf{r}(t_0)}{t-t_0}| = |\mathbf{r}'(t_0)|$.

Definition 1.18 *Let* $\mathbf{r} : I \to \mathbf{R}^n$ *denote a smooth vector function. At* $t_0 \in int(I)$, *it has* speed $v(t_0) = \mathbf{r}'(t_0)$ *and* unit tangent vector $\mathbf{t}(t_0) = \frac{\mathbf{r}'(t_0)}{|\mathbf{r}'(t_0)|}$. *In particular,*

$$\mathbf{r}'(t_0) = v(t_0)\mathbf{t}(t_0) \textit{ for all } t_0 \in int(I).$$

Illustration 1.19 *Parametrization: Interval, derivatives, tangents and speed curve.*

Reparametrizations

How are different parametrizations of the same regular curve related to each other? Well, they should give rise to the same tangent lines along the curve, whereas speed functions can be entirely different. Here is a method to *reparameterise* a regular curve with a given regular parametrization $\mathbf{r} : I \to \mathbf{R}^n$: Choose a (strictly monotone) smooth bijective[1] function $f : J \to I$ from an interval J such that $f'(t) > 0$ for all $t \in int(J)$. Then the composite vector function $\mathbf{r}_1 = \mathbf{r} \circ f : J \to \mathbf{R}^n$ parametrizes C – since $\mathbf{r}_1(t) = \mathbf{r}(f(t))$ or $\mathbf{r}(u) = \mathbf{r}_1(f^{-1}(u))$, and it is regular (Why? The chain rule in Prop. 1.9.5.)

Illustration 1.20 *Two intervals, a reparametrization, and the motion along the curve. (three moving spots)*

We use this approach to make the definition of a regular curve from Def. 1.13 more precise. Remark first that we may define an *equivalence relation* on regular parametrisations: $\mathbf{r} \simeq \mathbf{r}_1$ if and only if $\mathbf{r}_1$ is a reparametrization of $\mathbf{r}$; cf. Exc. 1.2.3.?.

[1] 1-1 and onto

Definition 1.21 *A* regular curve *is an equivalence class of regular parametrisations under the equivalence class $\simeq$ above.*

Since we only allow reparametrization functions f with $f' > 0$ everywhere, we can speak about an orientation (order) on the curve: $\mathbf{r}(t_0)$ comes before $\mathbf{r}(t_1)$ if and only if $t_0 < t_1$.

Arc Length

Let $\mathbf{r} : I \to \mathbf{R}^n$ denote a smooth vector function that is a regular parameterisation of a regular curve $C \subset \mathbf{R}^n$.

Definition 1.22 *The* length $l(C)$ *of the curve C is defined as the non-negative real number $\int_I |\mathbf{r}'(t)|\, dt$.*
For $\mathbf{r}(t) = [x_1(t), \ldots, x_n(t)]$, we obtain:

$$l(C) = \int_I \sqrt{(x_1'(t))^2 + \cdots + (x_n'(t))^2}\, dt.$$

Example 1.23

1. *For the helix from Ex. 1.4 with parametrization $\mathbf{r} : \mathbf{R} \to \mathbf{R}^3$, $\mathbf{r}(t) = [a\cos t, a\sin t, bt]$, we calculate:*

$$|\mathbf{r}'(t)| = \sqrt{(a\sin t)^2 + (a\cos t)^2 + b^2} = \sqrt{a^2 + b^2}.$$

   Hence, one turn around the helix has length

$$l = s(2\pi) = \int_0^{2\pi} \sqrt{a^2 + b^2} dt = 2\pi\sqrt{a^2 + b^2}.$$

   Remark that you obtain the familiar formula for the arc length of a circle in case $b = 0$.

2. *For the graph of a function $f : [a, b] \to \mathbf{R}$ from Ex. 1.7 with parametrization $\mathbf{r} : [a, b] \to \mathbf{R}^2$, $\mathbf{r}(t) = [t, f(t)]$, $a \le t \le b$, we calculate the length l of the graph as follows:*
   $|\mathbf{r}'(t)| = |[1, f'(t)]| = \sqrt{1 + (f')^2(t)}$, *and hence* $l = \int_a^b \sqrt{1 + (f')^2(t)} dt$.

The last example shows, that it is very often difficult or impossible to calculate the length of a curve in explicit terms. Many of the integrands involving square roots do not have explicit antiderivatives!

Here is a *motivation* for Def. 1.22: For $t \in \bar{I}$, let $I_t = I \cap]-\infty, t]$ denote the part of the interval containing numbers less than or equal to t. Then, the vector function $\mathbf{r}_t : I_t \to \mathbf{R}^n,\ \mathbf{r}_t(u) = \mathbf{r}(u)$ is a regular parametrization of a section C_t of the curve C. Let $s(t)$ denote the length of C_t. The function $s : I \to \mathbf{R}$ is called the *arc length function* of the parametrization. Note that s is monotonous: Increasing t yields increasing $s(t)$!

Arc length $s(t)$ and speed $v(t)$ of the parametrization $\mathbf{r}$ should be related by the property: $s'(t) = v(t) = |\mathbf{r}'(t)|$, and hence $s(t) = \int_I |\mathbf{r}'(t)|\, dt + c$, c a real constant. For $t_0 = \inf I$, the length of the empty curve C_{t_0} has to be zero, hence $c = 0$.

Illustration 1.24 *Scrollable interval, parametrization, section C_t, functions $s(t)$ and $s'(t)$.*

If the concept of length has any value, it should be independent of the choice of a regular parametrization of a given curve. This is indeed the case; see Exc. 1.2.3.6.

Unit Speed Parametrization

It is somehow awkward to have too many parametrizations for the same curve. For theoretical purposes, one needs a particular nice one:

Definition 1.25 *A vector function $\mathbf{r}_{al} : I \to \mathbf{R}^n$ is called a* unit speed *parametrization if its speed satisfies $v(s) = |\mathbf{r}'_{al}(s)| = 1$ for every $s \in int(I)$.*

In other words, the derivative of $\mathbf{r}_{al}$ is the *unit* tangent vector $\mathbf{t}(s)$ at every element $s \in I :\ \mathbf{t}(s) = \mathbf{r}'_{al}(s)$. It has the following property which explains its second name – an *arc length parametrization*:

Proposition 1.26 *For $t_0 \le t_1 \in I$, the arc length function s associated with a unit speed parametrization has the property: $s(t_1) - s(t_0) = t_1 - t_0$.*

Proof: $s(t_1) - s(t_0) = \int_{I_{t_1}} |\mathbf{r}'_{al}(s)|\, ds - \int_{I_{t_1}} |\mathbf{r}'_{al}(s)|\, ds = \int_{t_0}^{t_1} |\mathbf{r}'_{al}(s)|\, ds = \int_{t_0}^{t_1} ds = t_1 - t_0$.

□

Example 1.27 *The parametrization* $\mathbf{r}(t) = [a\cos t, a\sin t, bt]$ *for the helix from Ex. 1.4 yields (cf. Ex. 1.23):* $v(t) = |\mathbf{r}'(t)| = \sqrt{a^2+b^2}$*; in particular, it is a* constant *function. Hence, starting at* $t = 0$*, the arc length function is given as* $s(t) = \sqrt{a^2+b^2}t$*; and thus its inverse is* $t(s) = \frac{s}{\sqrt{a^2+b^2}}$*. Substituting* $t(s)$ *into the parametrization* $\mathbf{r}$ *yields the arc length parametrization for the helix, to wit:*

$$\mathbf{r}_{al}(s) = \mathbf{r}(t(s)) = [a\cos\frac{s}{\sqrt{a^2+b^2}}, a\sin\frac{s}{\sqrt{a^2+b^2}}, b\frac{s}{\sqrt{a^2+b^2}}].$$

Remark 1.28

1. *The following gives an easy* intuitive *idea for the arc length parametrization of a given curve: Imagine a piece of rope with a scale (starting at 0) and bend it along the curve (without stretching!) Then,* $\mathbf{r}_{al}(s)$ *is the vector from the origin to the point corresponding to the mark* s *on the scale.*

2. *The advantage of the arc length parametrization is that it focusses on the* geometric properties *of the curve rather than the infinitely many possible different modes (with varying speed, acceleration etc.) to run through it. Using this parametrization, it will be much easier to* define *entities like the curvature; on the other hand, for concrete* calculations*, one usually does not dispose of a concrete arc length parametrization.*

Does a regular curve always admit a unit speed parametrization? Well, in most cases, it is not possible to write down an explicit unit speed parametrization using familiar functions as components. Nevertheless, a unit speed parametrization always exists:

To find it, let us first assume that there is a unit speed parametrization $\mathbf{r}_{al} : I \to \mathbf{R}^n$ representing a curve C. Let $f : J \to I$ denote any bijection with $f'(t) > 0$ giving rise to the (re)parametrization $\mathbf{r} = \mathbf{r}_{al} \circ f$ of the curve C. The two parametrizations are linked via the arc length function $s : J \to [0, l(C)]$ associated to the parametrization $\mathbf{r}$, to wit:

$$\mathbf{r}(t) = \mathbf{r}_{al}(s(t)).$$

("At time t, we trace the point at distance $s(t)$ from t he start point"). How could one (re)construct $\mathbf{r}_{al}$ given $\mathbf{r}$? Well, if the function s had an inverse $t : [0, l(C)] \to J$, then

$$\mathbf{r}_{al}(s) = \mathbf{r}_{al}(s(t(s))) = \mathbf{r}(t(s)). \tag{1.1}$$

Illustration 1.29 *Scrollable* s *and* t*-intervals and parametrization.*

The function $s : J \to \mathbf{R},\ s(t) = \int_I |\mathbf{r}'(t)\, dt$ has the derivative $s'(t) = |\mathbf{r}'(t)| = f'(t) > 0$, and hence s is strictly increasing. Therefore, it has in fact a smooth inverse, and we can use (1.1) as the *definition* of an arc length (unit speed) parametrization. Remark that we have shown the existence of a unit speed parametrization using the existence of a function t inverse to s. In most cases, it is not possible to give a formula for $t(s)$ in terms of well-known functions. Just try for the graph of a differentiable function (cf. Ex. 1.7).

As you may have observed by now, it is customary to use the parameter s for an arc length parameter, and t for an arbitrary parameter. Whenever s is used in this text, we talk about a unit speed parametrization.

A geometric property of unit speed parametrizations

The following property of unit speed parametrizations is the main reason why geometers like them so much. It is an immediate consequence of the fundamental trick (Prop. 1.10):

Proposition 1.30 *Let $\mathbf{r} : I \to \mathbf{R}^n$ denote a unit speed parametrization. Then,*

$$\mathbf{r}''(s) \cdot \mathbf{r}'(s) \ = \ 0 \textit{ for all } s \in int(I).$$

In geometric terms, the acceleration vector $\mathbf{r}''(s)$ is always perpendicular on the tangent vector $\mathbf{r}'(s)$.

1.2.3 Exercises

1. Calculate tangents of several curves and use applet to show them.

2. Determine a parametrization for the tangent line to the graph of the function $f\ :]a, b[\to \mathbf{R}$ (with parametrization $\mathbf{r}\ :]a, b[\to \mathbf{R}^2$, $\mathbf{r}(t) = [t, f(t)]$ at $t_0 \in]a, b[$. Is the result familiar?

3. How can you determine from $\mathbf{r}'(t_0)$ whether the curve C has a horizontal/vertical tangent at P_{t_0}?

4. Why can the curve from Ex. 1.11 not have any regular parametrization at $P_0 : [1, 1]$?

5. Show that the relation $\simeq$ on regular parametrisations ($\mathbf{r} \simeq \mathbf{r}_1$ if and only if there is a bijective function $f : J \to I$ with $f'(t) > 0$ for every $t \in int(J)$

with $\mathbf{r}_1 = \mathbf{r} \circ f$) is an equivalence relation (i.e., reflexive, symmetric, and transitive).

One may change the definition of $\simeq$ by allowing bijective reparametrisations f with $f'(t) \neq 0$ for every $t \in int(J)$. How is the geometric meaning changed?

6. Show that the length of a regular curve C is independent of the choice of a regular parametrization.
(Hint: Let $\mathbf{r} : I \to \mathbf{R}^n$ denote a regular parametrization of C, let $f : J \to I$ denote a real function such that $\mathbf{r}_1 = \mathbf{r} \circ f : J \to \mathbf{R}^n$ is a regular reparametrization of C. Using the chain rule for derivatives and substitution for integals, show, that

$$\int_J |\mathbf{r}_1'| \, du = \int_I |\mathbf{r}'| \, dt.)$$

7. Find a unit speed parametrization of the regular space curve C given by the parametrization $\mathbf{r}(t) = [t, \frac{\sqrt{6}}{2}t^2, t^3],\ t \in \mathbf{R}$.

1.3 Curvature

From now on, we restrict ourselves to (regular) plane and space curves, given by regular smooth parametrizations $\mathbf{r} : I \to \mathbf{R}^n,\ n = 2$ or $n = 3$.

What should the curvature of a curve be? Well, a line is not curved at all; its curvature has to be zero. A circle with a small radius is more "curved" than a circle with a large radius. Circles and lines have *constant* curvature. Curves that are not (pieces of) circles or lines will have a curvature *varying* from point to point.

1.3.1 Approximating and osculating circles

A tangent line to a curve C can be defined as the limit of the secant lines through *two* points on the curve. In much the same spirit, one can define curvature circles (= osculating circles) as the limit circles of circles passing through *three* points on the curve. The curvature of C at a point P is then defined as the inverse of the radius of this curvature circle at P.

Let C denote a plane or space curve given by a *unit speed* parametrization $\mathbf{r} : I \to \mathbf{R}^n,\ n = 2$ or $n = 3$. Let $s_0 \in int(I)$ and P_{s_0} the point on C with $\mathbf{r}(s_0) = \overrightarrow{OP_{s_0}}$.

To construct the curvature circle at P_{s_0}, choose s_1 and s_2 close to s_0 such that P_{s_0}, P_{s_1} and P_{s_2} are not contained in a line (This is possible unless the curve is coinciding with a line close to P_s). In Exc. 1.3.7.?, you are invited to show that there exists exactly one circle through the points P_{s_0}, P_{s_1} and P_{s_2}; we denote its center by $C(s_0, s_1, s_2)$.

The idea is to find a limit circle when both s_1 and s_2 approach s_0. To this end, we investigate the real smooth function

$$d(s) = (\mathbf{r}(s) - C(s_0, s_1, s_2)) \cdot (\mathbf{r}(s) - C(s_0, s_1, s_2)).$$

Its geometric meaning is the square of the distance between P_s and $C(s_0, s_1, s_2)$. By definition, $d(s_0) = d(s_1) = d(s_2)$, i.e., the square of the radius of the circle through these three points. By Rolle's theorem for smooth functions in one variable, there exist two intermediate values t_1, t_2 such that $d'(t_1) = d'(t_2)$, and, moreover an intermediate value u such that $d''(u) = 0$. The derivatives of d can be calculated as

$$\begin{aligned} d'(s) &= 2\mathbf{r}'(s) \cdot (\mathbf{r}(s) - C(s_0, s_1, s_2)), \text{ resp.} \qquad (1.2) \\ d''(s) &= 2(\mathbf{r}''(s) \cdot (\mathbf{r}(s) - C(s_0, s_1, s_2)) + 2\mathbf{r}'(s) \cdot \mathbf{r}'(s) \\ &= 2(\mathbf{r}''(s) \cdot (\mathbf{r}(s) - C(s_0, s_1, s_2)) + 2. \qquad (1.3) \end{aligned}$$

To obtain the last equality, we used that r is a unit speed parametrization. At $s = u$, we obtain in particular: $\mathbf{r}''(u) \cdot (\mathbf{r}(u) - C(s_0, s_1, s_2)) = -1$.

Illustration 1.31 *Circles through 3 points on the curve converging to osculating circle. The function d and intermediate points. See geometric lab.*

Now, let s_1 and s_2 tend to s_0 to get a limiting curvature circle with centre $C(s_0)$ and radius $\rho(s_0)$. You are invited to show the existence of this curvature circle in Exc. 1.3.7.1. When s_1 and s_2 tend to s, t_1, t_2 and u tend to s_0, as well. This allows us to derive from (1.2), that $\mathbf{r}'(s_0) \cdot (\mathbf{r}(s_0) - C(s_0)) = 0$, and that $\mathbf{r}''(s_0) \cdot (\mathbf{r}(s_0) - C(s_0)) = -1$.

We obtain the following geometric implications:

Lemma 1.32
- *The tangent line to C at P_{s_0} (parallel to $\mathbf{t}(s_0) = \mathbf{r}'(s_0)$) is perpendicular to the radial vector $\overrightarrow{P_{s_0}C_{s_0}}$ and, for a* plane *curve*
- *The acceleration vector $\mathbf{t}'(s_0) = \mathbf{r}''(s_0)$ is perpendicular to $\mathbf{r}'(s_0)$ according to Prop. 1.30, and hence, it has to be* parallel *to $\mathbf{r}(s_0) - C(s_0)$; calculating lengths, we obtain: $|\mathbf{r}''(s_0)|\,|\mathbf{r}(s_0) - C(s_0)| = 1$, i.e., $|\mathbf{t}'(s_0)| = |\mathbf{r}''(s_0)| = \frac{1}{\rho(s_0)}$. The length of the acceleration vector is thus inverse to the radius of curvature at s_0.*

This length $|\mathbf{t}'(s_0)| = |\mathbf{r}''(s_0)|$ is the absolute value of the curvature of the plane curve C at P_{s_0}. We shall introduce a refinement, the signed curvature for plane curves in the next Sect. 1.3.2.

Illustration 1.33 *Curve,* $\mathbf{r}'(s_0), \mathbf{r}(s_0) - C(s_0), \mathbf{r}''(s_0), |\mathbf{r}''(s_0)|$. *Scrollable.*

For a space curve C, one can follow the same line of argument; we need the existence of the osculating plane and its properties to complete the description of the curvature for space curves. See Sect. 1.4.1.

1.3.2 Principal normals and curvature functions

Let C be a regular smooth curve in the plane or in 3-space with unit speed parametrization $\mathbf{r} : I \to \mathbf{R}^n$, $n = 2$ or 3. By Def. 1.25 – and its explanation, the vector $\mathbf{t}(s) = \mathbf{r}'(s)$ is a unit tangent vector to the curve at the point P_s. The vector function $\mathbf{t} : I \to \mathbf{R}^i$ is called the *unit tangent vector field* moving along the curve. We are now going to analyse the information hidden in the derived vector field $\mathbf{t}' = \mathbf{r}''$ along the curve C. An application of Prop. 1.30 yields:

Corollary 1.34 *1. At every point P_s of the curve, the derivative $\mathbf{t}'(s)$ is* perpendicular *to $\mathbf{t}(s)$: $\mathbf{t}'(s) \cdot \mathbf{t}(s) = 0$.*

2. For a plane *curve C the vectors $\mathbf{t}'(s)$ and $\hat{\mathbf{t}}(s)$ are* parallel.

Definition 1.35 *1. The vector $\mathbf{t}'(s)$ is called the* curvature vector *at the point P_s on C with $\overrightarrow{OP_s} = \mathbf{r}(s)$.*

2. A point P_s on C is called an inflection point *if $\mathbf{t}'(s) = \mathbf{0}$.*

3. The principal normal vector *$\mathbf{n}(s)$ to the curve at P_s is defined as follows:*

(a) For a plane *curve C let $\mathbf{n}(s) = \hat{\mathbf{t}}(s)$.*

(b) For a space *curve C let $\mathbf{n}(s) = \frac{\mathbf{t}'(s)}{|\mathbf{t}'(s)|}$ at every non-inflection point $P_s \in C$.*

Definition 1.36 *The* curvature *$\kappa(P_s)$ of the curve C at the point P_s is defined as follows:*

1. For a plane *curve C let*

$$\mathbf{t}'(s) = \kappa(s)\mathbf{n}(s), \tag{1.4}$$

and thus $\kappa(s) = \pm|\mathbf{t}'(s)|$.

2. *For a* space *curve* C *let*

$$\kappa(s) = |\mathbf{t}'(s)| \geq 0. \tag{1.5}$$

and thus for $\kappa(s) \neq 0$: $\mathbf{t}'(s) = \kappa(s)\mathbf{n}(s)$.

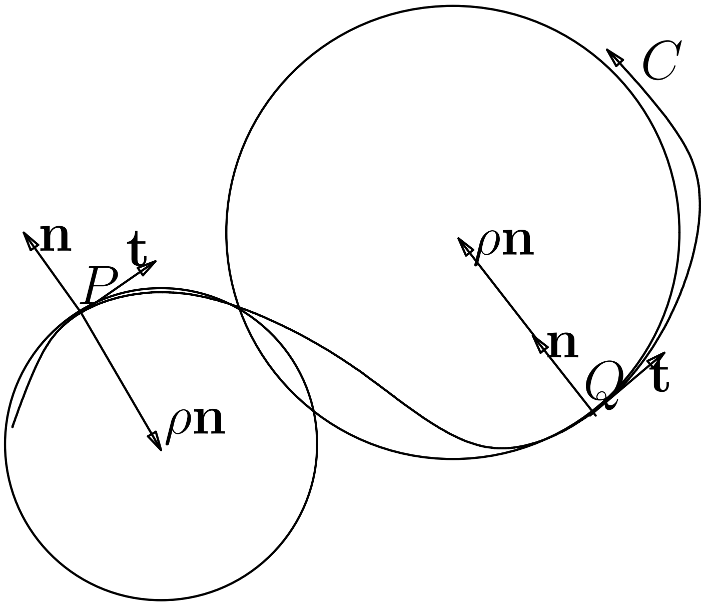


Figure 1.3: Tangent and principal normal vectors and osculating circles at points P and Q

Remark 1.37 *1. By Cor. 1.34, tangent and principal normal vectors are perpendicular to each other at every point of the curve:* $\mathbf{t}(s) \cdot \mathbf{N}(s) = 0$.

2. An inflection point P_s *on* C *is characterised by the property* $\kappa(s) = 0$.

Illustration 1.38 *Moving tangent,* $\mathbf{t}'(s), \mathbf{n}(s)$*, osculating circle; view inflection points; moreover, plot curvature.*

Here are several *motivations* for the definition of the curvature above:

1. The *magnitude* $|\mathbf{t}'(s)|$ of the derivative $\mathbf{t}'(s)$ measures the *rate of change* of the *direction* of the tangent vector field $\mathbf{t}(s)$ – since its length $|\mathbf{t}'(s)|$ is constant. The faster the direction of the tangent vector changes, the more curved is the curve.

2. Let us calculate the curvature for a plane circle with radius R. With center at the origin, the arc length parametrization of the circle is given by $\mathbf{r}_{al}(s) = [R\cos\frac{s}{R}, R\sin\frac{s}{R}]$. We calculate:

$$\mathbf{t}(s) = [-\sin\frac{s}{R}, \cos\frac{s}{R}],\ \mathbf{t}'(s) = [-\frac{1}{R}\cos\frac{s}{R}, -\frac{1}{R}\sin\frac{s}{R}] = \frac{1}{R}\hat{\mathbf{t}}(s),$$

   and according to (1.4), $\kappa(s) = \frac{1}{R}$ for every s. Hence, the curvature of a circle is constant and in inverse proportion to its radius – as it should!

   For a general curve C, the curvature at a point P is in inverse proportion to the radius of the *best approximating circle* at P, the so-called osculating circle, cf. Fig. 1.3 and Sect. 1.3.1.

1.3.3 Tangent and normal indicatrix and curvature

In the plane, a unit vector field $\mathbf{v} : I \to \mathbf{R}^2$ takes values in the unit circle $S^1 := \{[x, y] \in \mathbf{R}^2 |\, x^2 + y^2 = 1\} = \{[\cos(\theta), \sin(\theta)] |\, \theta \in \mathbf{R}\}$. It can thus be seen as a continuous map $\mathbf{v} : I \to S^1$; if the curve does not have double points[2], one may view it as a map from C to S^1, the so-called *indicatrix*. We shall later use a similar point of view to define the Gauss map on a surface.

Both the unit tangent vector field $\mathbf{t}(s)$ and the principal normal vector field $\hat{\mathbf{t}}(s)$ give rise to (perpendicular) *tangent* and *normal indicatrices.*

Illustration 1.39 *Tangent and normal indicatrices of a plane curve; simultaneously moving points on interval, curve and S^1.*

To relate to curvature calculations, we choose a description of the unit tangent vector field as $\mathbf{t}(s) = [\cos\theta(s), \sin\theta(s)]$ with $\theta(s)$ the angle between $\mathbf{t}(s)$ and the horizontal vector $\mathbf{i}$, i.e., $cos\theta(s) = \mathbf{t}(s)\cdot\mathbf{i}$[3]. A calculation of $\mathbf{t}'(s)$ using the chain rule (Prop. 1.9(5)) yields:

$$\mathbf{t}'(s) = \theta'(s)[-\sin\theta(s), \cos\theta(s)] = \theta'(s)\hat{\mathbf{t}}(s), \tag{1.6}$$

and hence: $\kappa(s) = \theta'(s)$. Hence, the curvature measures the *rate of change* for the *angle* between tangents, as it should. Moreover, we get an explananation for the *sign* of the curvature of a plane curve, to wit:

Corollary 1.40 *Near the point P_s, a plane curve C is curved*

[2] and neither converges to a point already on the curve at an end of the interval

[3] This can easily be done locally and then "lifted" along the entire curve. Try Exc. 1.3.7.

- counter-clockwise *if and only if* $\kappa(s) > 0$,
- *and* clockwise *if and only if* $\kappa(s) < 0$.

In Fig. 1.3, the curvature is negative at P and positive at Q.

Illustration 1.41 *Banchoff's curvature caterpillar. One needs to explain, that* $\mathbf{n}'(s) = -\kappa(s)\mathbf{t}(s)$ *– Exc. 1.3.7.?!*

Indicatrices for space curves will be covered in Sect. 1.4.1.

1.3.4 Acceleration and curvature calculation

The definition of curvature above uses the unit speed parametrization of a given curve. But in general, you have only a *regular* parametrization $\mathbf{r} : J \to \mathbf{R}^n$ at hand. In this case, the curvature at a given point P is "hidden" in the *acceleration vector* at that point. Let us first compare the derivatives of different parametrizations of the *same* curve C:

Let $\mathbf{r} : J \to \mathbf{R}^n$ denote a regular smooth parametrization, let $s : J \to I = s(J) \subset \mathbf{R}$ denote the associated arc length function $s(t) = \int_I^t |\mathbf{r}'(u)|\, du$. Let $\mathbf{r}_{al} : I \to \mathbf{R}^n$ the unit speed (arc length) reparametrization (cf. Sect. 1) with the property

$$\mathbf{r}(t) = \mathbf{r}_{al}(s(t)).$$

Let us calculate the derivatives of the parametrization $\mathbf{r}(t)$ using the chain rule and the product rule from Prop. 1.9 and the definition of curvature from Def. 1.36:

$$\mathbf{r}'(t) = s'(t)\mathbf{r}'_{al}(s(t)) = v(t)\mathbf{t}(t);$$

$$\mathbf{r}''(t) = v'(t)\mathbf{r}'_{al}(s(t)) + (s'(t))^2\mathbf{r}''_{al}(s(t)) = v'(t)\mathbf{t}(t) + v(t)^2\kappa(t)\mathbf{n}(t). \tag{1.7}$$

(If $P_t \in C \subset \mathbf{R}^n$ is the point with $\overrightarrow{OP_t} = \mathbf{r}(t)$, then $v(t)$ is the speed at P_t associated to the parametrization, $a(t) = v'(t) = s''(t)$ is the *scalar acceleration* at P_t, $\mathbf{t}(t)$ is the unit tangent vector to C at P_t, $\kappa(t)$ the curvature of C and $\mathbf{n}(t)$ the principal normal vector to C at P_t.)

Before using (1.7) to calculate the curvature of a given curve, let us look at the following attractive *interpretation in mechanics*: Equation (1.7) yields a decomposition of the *acceleration* vector $\mathbf{a}(t) = \mathbf{r}''(t)$ into a *tangential* component $\mathbf{a_t}(t)$ and a *normal* component $\mathbf{a_n}(t)$:

$$\mathbf{a}(t) = \mathbf{a_t}(t) + \mathbf{a_n}(t) = v'(t)\mathbf{t}(t) + v^2(t)\kappa(t)\mathbf{n}(t).$$

In particular, the magnitude of the *tangential* component is: $|\mathbf{a_t}(t)| = v'(t)$, which is the *scalar accelaration*, i.e., the rate of change of the speed. The magnitude of the *normal* component is: $|\mathbf{a_n}(t)| = v^2(t)\kappa(t)$. Hence, the force acted upon a particle *normal to its path* is proportional to the *square of its speed* and to the *curvature of the curve.* This is intuitively known to every car driver; when you drive through a narrow curve, you have to slow down drastically in order to avoid strong normal forces.

Curvature formulas

Proposition 1.42 *1. Let C be a* plane *curve with parametrization* $\mathbf{r} : J \to \mathbf{R}^2$*. Its curvature $\kappa(t)$ at a point P_t with $\overrightarrow{OP_t} = \mathbf{r}(t)$ is given by*

$$\kappa(t) = \frac{[\mathbf{r}'(t), \mathbf{r}''(t)]}{|\mathbf{r}'(t)|^3}. \tag{1.8}$$

The numerator $[\mathbf{r}'(t), \mathbf{r}''(t)]$ *is the* plane product *of the vectors* $\mathbf{r}(t), \mathbf{r}'(t) \in \mathbf{R}^2$*.*

More explicitly, for $\mathbf{r}(t) = [x(t), y(t)]$, $t \in J$, *we obtain:*

$$\kappa(t) = \frac{\begin{vmatrix} x'(t) & x''(t) \\ y'(t) & y''(t) \end{vmatrix}}{(\sqrt{x'(t)^2 + y'(t)^2})^3}. \tag{1.9}$$

2. Let C be a space *curve with parametrization* $\mathbf{r} : J \to \mathbf{R}^3$*. Its curvature $\kappa(t)$ at a point P_t with $\overrightarrow{OP_t} = \mathbf{r}(t)$ is given by*

$$\kappa(t) = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|^3}. \tag{1.10}$$

More explicitly, for $\mathbf{r}(t) = [x(t), y(t), z(t)]$, $t \in J$, *we obtain:*

$$\kappa(t) = \frac{|[x'(t), y'(t), z'(t)] \times [x''(t), y''(t), z''(t)]|}{(\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2})^3}. \tag{1.11}$$

Proof:

1. Using properties of the determinant of a (2×2)-matrix and (1.7), we obtain: $[\mathbf{r}'(t), \mathbf{r}''(t)] = [v(t)\mathbf{t}(t), v'(t)\mathbf{t}(t) + v^2(t)\kappa(t)\hat{\mathbf{t}}(t)] = [v(t)\mathbf{t}(t), v^2(t)\kappa(t)\mathbf{n}(t)] = v^3(t)\kappa(t)[\mathbf{t}(t), \hat{\mathbf{t}}(t)] = v^3(t)\kappa(t)$.
(The last equation uses that the plane product of a unit vector and its hat vector is one, since it measures the (signed) area of the rectangle spanned by both vectors.)

2. Using properties of the cross product and (1.7), we obtain:
$\mathbf{r}'(t)\times\mathbf{r}''(t) = v(t)\mathbf{t}(t)\times(v'(t)\mathbf{t}(t)+v^2(t)\kappa(t)\mathbf{n}(t)) = v(t)\mathbf{t}(t)\times v^2(t)\kappa(t)\mathbf{n}(t) = v^3(t)\kappa(t)(\mathbf{t}(t)\times\mathbf{n}(t))$.
The vector $\mathbf{t}(t)\times\mathbf{n}(t)$ is a unit vector, since $\mathbf{t}(t)$ and $\mathbf{n}(t)$ are mutually orthogonal unit vectors. Therefore, the vector $\mathbf{r}'(t)\times\mathbf{r}''(t)$ has length $v^3(t)\kappa(t)$.

□

Example 1.43 *1. An ellipse E with semi-axes a and b has a parametrization $\mathbf{r}(t) = [a\cos t, b\sin t]$, $t\in[0,2\pi]$. We calculate:*

$$\begin{aligned}\mathbf{r}'(t) &= [-a\sin t, b\cos t],\\ \mathbf{r}''(t) &= [-a\cos t, -b\sin t],\end{aligned}$$

$$[\mathbf{r}'(t),\mathbf{r}''(t)] = \begin{vmatrix} -a\sin t & -a\cos t\\ b\cos t & -b\sin t\end{vmatrix} = ab.$$

Since $|\mathbf{r}'(t)| = \sqrt{a^2(\sin t)^2+b^2(\cos t)^2}$, we obtain:

$$\kappa(t) = \frac{ab}{(a^2(\sin t)^2+b^2(\cos t)^2)^{\frac{3}{2}}}. \tag{1.12}$$

2. The curve C_f given as the graph *of a function $f:[a,b]\to\mathbf{R}$ (Ex. 1.7) can be parameterised by the vector function $\mathbf{r}:[a,b]\to\mathbf{R}^2$ given as $\mathbf{r}(t)=[t,f(t)]$. We calculate: $\mathbf{r}'(t)=[1,f'(t)]$, $\mathbf{r}''(t)=[0,f''(t)]$; hence, $[\mathbf{r}'(t),\mathbf{r}''(t)] = f''(t)$, $|\mathbf{r}'(t)| = \sqrt{1+(f'(t))^2}$, and*

$$\kappa(P_t) = \frac{f''(t)}{\sqrt{1+(f'(t))^2}^3}. \tag{1.13}$$

In particular, C_f is curved counter-clockwise at P_t if $f''(t)>0$ and clockwise if $f''(t)<0$ (Cor. 1.40).

3. Let $\mathbf{r}$ denote the parametrization for a helix from Ex. 1.4 given as

$$\mathbf{r}(t) = [a\cos t, a\sin t, bt], a, b>0.$$

Its derivatives (velocity and acceleration vectors) are calculated as

$$\begin{aligned}\mathbf{r}'(t) &= [-a\sin t, a\cos t, b];\\ \mathbf{r}''(t) &= [-a\cos t, -a\sin t, 0].\end{aligned} \tag{1.14}$$

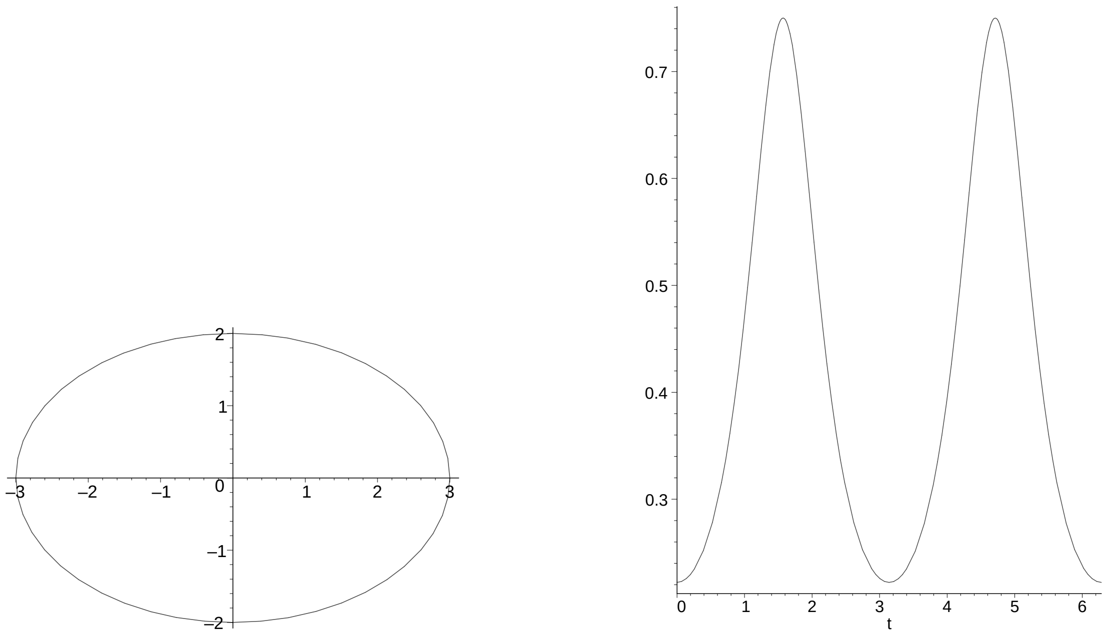


Figure 1.4: Ellipse and curvature function

Hence,

$$\begin{aligned} |\mathbf{r}'(t)| &= \sqrt{a^2+b^2}; \\ \mathbf{r}'(t)\times\mathbf{r}''(t) &= [ab\sin t, -ab\cos t, a^2]; \\ |\mathbf{r}'(t)\times\mathbf{r}''(t)| &= a\sqrt{a^2+b^2}. \end{aligned} \tag{1.15}$$

The curvature of the helix at P_t is calculated as

$$\kappa(t) = \frac{a\sqrt{a^2+b^2}}{\sqrt{a^2+b^2}^3} = \frac{a}{a^2+b^2}. \tag{1.16}$$

Note, that the curvature is constant *along the helix. For $b = 0$, we get as a special case the curvature $\frac{1}{a}$ of a circle with radius a.*

Osculating Circles

In Sect. 1.3.1, we motivated the curvature concept via the radii of osculating circles approximating the curve. Having determined the curvature of a curve with a given parametrization, we can now determine the osculating circles for a plane curve C given by a regular smooth parametrization $\mathbf{r} : I \to \mathbf{R}^2$.

Proposition 1.44 *The osculating circle at P_t with $\overrightarrow{OP_t} = \mathbf{r}(t)$ has radius $\rho(t) = \frac{1}{|\kappa(t)|}$ and center C_t with $\overrightarrow{P_tC_t} = \frac{1}{\kappa(t)}\mathbf{n}(t)$.*

Proof: Exc. 1.3.7.?. Use Lemma 1.32.

□

In a similar way, one can determine the osculating circles along a space curve. In that case, one has to use that the osculating circle through P_t is contained in the osculating plane $\omega(t)$ through P_t (cf. Sect. 1.4.1).

The Evolute Curve

From a given curve C, one may obtain a new curve E_C, the *evolute* of C, by associating to every point $P \in C$ the corresponding *centre of curvature* C_P, i.e., the center of the osculating circle through P. It is easy to translate the description of this center in 1.44 into a parametrization of E_C:

Corollary 1.45 *Let C be a curve with parametrization $\mathbf{r} : I \to \mathbf{R}^2$. The following is a parametrization for the evolute E_C of C:*

$$\mathbf{e}(t) = \mathbf{r}(t) + \frac{\mathbf{n}(t)}{\kappa(t)} = \mathbf{r}(t) + \frac{|\mathbf{r}'(t)|^3}{[\mathbf{r}'(t), \mathbf{r}''(t)]}\frac{\widehat{\mathbf{r}'}(t)}{|\mathbf{r}'(t)|} = \mathbf{r}(t) + \frac{|\mathbf{r}'(t)|^2}{[\mathbf{r}'(t), \mathbf{r}''(t)]}\widehat{\mathbf{r}'}(t).$$

Example 1.46
1. *The evolute of a cycloid is a translated cycloid.*
2. *The evolute of an ellipse is an* astroid. *Using Cor. 1.45, we obtain the parametrization $\mathbf{e}(t) = [\frac{a^2-b^2}{a}(\cos t)^3, \frac{b^2-a^2}{b}(\sin t)^3]$.*

Illustration 1.47 *Replace figure by an applet producing moving osculating circles and the evolute (as done by Robert for a specific curve)*

Proposition 1.48 *Let C denote a regular plane curve with* constant *curvature κ. For $\kappa = 0$, C is a straight line segment; for $\kappa \neq 0$, C is contained in a circle.*

Proof: Let $\mathbf{r}$ denote a unit speed parametrization of C. If $\kappa = 0$, then $\mathbf{r}'' = \mathbf{0}$ (why?), and $\mathbf{r}'(s)$ is a constant vector $\mathbf{v}$. Hence, $\mathbf{r}(s) = s\mathbf{v} + \mathbf{b}$ for some constant vector $\mathbf{b}$; the image of an interval is thus contained in a straight line.

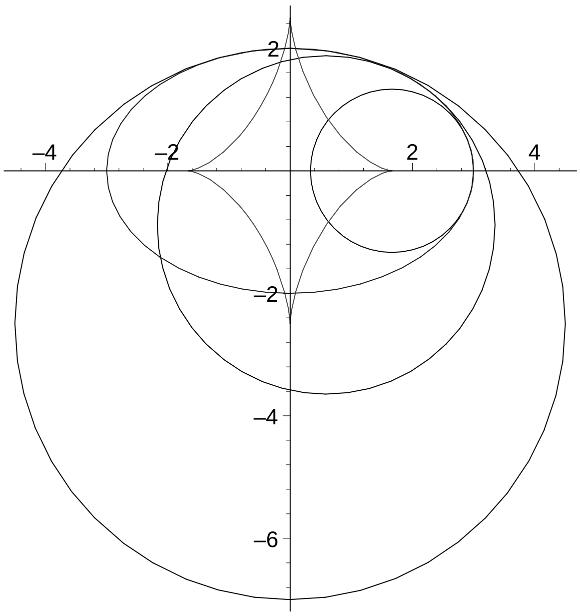


Figure 1.5: Ellipse with osculating circles and evolute

Now suppose $\kappa \neq 0$. The evolute curve of a circle is a constant point $\mathbf{e}(t) = \mathbf{e}$. Let us first show, that the evolute curve of C is constant, too: By Cor. 1.45, it is given by $\mathbf{e}(s) = \mathbf{r}(s) + \frac{\mathbf{n}(s)}{\kappa}$. Hence,

$$\mathbf{e}'(s) = \mathbf{r}'(s) + \frac{\mathbf{n}'(s)}{\kappa} = \mathbf{t}(s) - \frac{\kappa}{\kappa}\mathbf{t}(s) = \mathbf{0}.$$

(The middle equation stems from Exc. 1.3.7.?(c)). As a result, $\mathbf{e}(s)$ is a constant vector $\mathbf{e}$, and $\mathbf{r}(s) - \mathbf{e} = \frac{\mathbf{n}(s)}{\kappa}$ has constant length $|\mathbf{r}(s) - \mathbf{e}| = \frac{1}{|\kappa|}$. Hence, the image of $\mathbf{r}$ is contained in the circle with center at $\mathbf{e}$ and radius $\frac{1}{|\kappa|}$.

□

1.3.5 The Curvature Function characterizes a Plane Curve

We have seen how to associate to a curve its curvature function.

Illustration 1.49 *Again. Curve and asociated curvature function.*

Is it possible to somehow reverse this process? More precisely: Let $k : I \to \mathbf{R}$ denote a continuous function. Is there always a plane curve C (with unit speed parametrization $\mathbf{r} : I \to \mathbf{R}^2$) realizing this function as the curvature function associated to C? How many such curves are there?

The answer is, that there is always such a curve, and it is *uniquely determined up to a rigid motion* of the plane, i.e., up to a combination of a translation and a rotation. First of all, it is clear, that a translation, resp. a rotation does not change the curvature function κ. On the other hand the following iterative construction shows why one should expect the existence of a curve with given curvature function[4]:

Choose a start point P_0 and a start direction given by a unit vector $\mathbf{v}_0$ (this choice corresponds to the choice of a translation and a rotation). Draw a circular arc of length s through P_0 with tangent vector $\mathbf{v}_0$ and radius $\rho_0 = \frac{1}{\kappa(0)}$. At the end of this little circle you obtain a point P_s and a tangent direction $\mathbf{v}_s$. Continue with a circular arc of length s through P_s with tangent vector $\mathbf{v}_s$ and radius $\rho_s = \frac{1}{\kappa(s)}$ and obtain an end point P_{2s} and an end direction $\mathbf{v}_{2s}$. Keep on. The resulting curve will not be smooth everywhere, but for small s, the curvature will be a step function close to the original κ. Before the computer age, this method was in fact sometimes used to graph a curve given by a parametrization!

Illustration 1.50 *Curve consisting of circular arcs and curvature step function*

Let us now transform this intuitive reasoning into a theorem with a strict proof: Let $I \subset \mathbf{R}^2$ denote an open interval, let $C^k(I, \mathbf{R}^n)$ denote the continuous (vector) functions $\mathbf{r} : I \to \mathbf{R}^n$ that are k times differentiable on I with continuous derivatives[5].

Theorem 1.51 ***(The fundamental theorem for plane curves)*** *Let $k : I \to \mathbf{R}$ denote a smooth function. Then there is a curve with unit speed parametrization $\mathbf{r} : I \to \mathbf{R}^2$ with curvature function $\kappa(s) = k(s),\ s \in I$. Furthermore, any two such curves differ by a proper rigid motion.*

Proof:

Existence Let $k : I \to \mathbf{R}$ denote a smooth function and let $s_0 \in I$. Define

1. an angle function $\varphi : I \to \mathbf{R}$ as $\varphi(s) = \int_{s_0}^{s} k(u)\, du + \varphi_0$ for arbitrary $\varphi_0 \in \mathbf{R}$; this function satisfies $\varphi' = k$ by the fundamental theorem of calculus.

[4]Another way to introduce this construction is: How to find a curve whose curvature function is a step function

[5]Should we introduce this notation earlier?

2. a *unit* vector function $\mathbf{t} : I \to S^1 \subset \mathbf{R}^2$ by $\mathbf{t}(s) = [\cos\varphi(s), \sin\varphi(s)]$
3. a vector function $\mathbf{r} : I \to \mathbf{R}^2$ by
   $\mathbf{r}(s) = [\int_{s_0}^{s} \cos\varphi(u)\, du\, + x_0,\ \int_{s_0}^{s} \sin\varphi(u)\, du\, + y_0]$

To summarize, the parametrization $\mathbf{r}$ is essentially constructed by a *double integration* of the function k.

Let us check that the curvature function $\kappa : I \to \mathbf{R}$ associated to $\mathbf{r}$ agrees with the function k: By the fundamental theorem of calculus again, we have that $\mathbf{r}'(s) = \mathbf{t}(s),\ s \in I$. And $\mathbf{t}'(s) = \frac{d}{ds}[\cos\varphi(s), \sin\varphi(s)] = \varphi'(s)[-\sin\varphi(s), \cos\varphi(s)] = k(s)\hat{t}(s) = k(s)\mathbf{n}(s)$. Hence $\kappa(s) = k(s),\ s \in I$.

Uniqueness up to a rigid motion The equations

$$\begin{aligned} \mathbf{t}' &= \kappa\mathbf{n} \\ \mathbf{n}' &= -\kappa\mathbf{t} \end{aligned}$$

can be summarized in a matrix equation $\mathbf{F}' = \mathbf{F}\mathbf{A}$ with
$\mathbf{F}(s) = [\mathbf{t}(s)\mathbf{n}(s)] \in SO(2)$ and the skew-symmetric (!) matrix
$\mathbf{A}(s) = \begin{bmatrix} 0 & -\kappa(s) \\ \kappa(s) & 0 \end{bmatrix}$.

Let $\mathbf{r}_1, \mathbf{r}_2$ denote two parametrizations with $\kappa_1(s) = \kappa_2(s) =: \kappa(s)$. The associated matrices satisfy: $\mathbf{F}_1' = \mathbf{F}_1\mathbf{A}$ and $\mathbf{F}_2' = \mathbf{F}_2\mathbf{A}$ with the *s*ame matrix $\mathbf{A}(s)$ on the right hand side. Consider the matrix function $\mathbf{M}(s) = \mathbf{F}_1(s)\mathbf{F}_2^T(s) : I \to SO(2)$ and calculate its derivative with respect to s:

$$\begin{aligned} \mathbf{M}'(s) = \mathbf{F}_1'(s)\mathbf{F}_2^T(s) + \mathbf{F}_1(s)\mathbf{F}_2'^T(s) &= \mathbf{F}_1(s)\mathbf{A}(s)\mathbf{F}_2^T(s) + \mathbf{F}_1(s)\mathbf{A}^T(s)\mathbf{F}_2^T(s) \\ &= \mathbf{F}_1(s)(\mathbf{A}(s) + \mathbf{A}^T(s))\mathbf{F}_2^T(s) = 0 \end{aligned}$$

since $\mathbf{A}(s)$ is skew-symmetric. We conclude that $\mathbf{M}(s) = \mathbf{M}$ a constant matrix. In particular, $\mathbf{F}_1(s) = \mathbf{F}_1(s)\mathbf{F}_2^T(s)\mathbf{F}_2(s) = \mathbf{M}\mathbf{F}_2(s)$.

Now consider the curves given by $\mathbf{r}_1$ and $\mathbf{M}\mathbf{r}_2$ (after applying the rotation $\mathbf{M}$ to $\mathbf{r}_2$). Then $\frac{d}{ds}\mathbf{M}\mathbf{r}_2 = \mathbf{M}\mathbf{t}_2$ is the first column of the matrix $\mathbf{M}\mathbf{F}_2 = \mathbf{F}_1$, and hence $(\mathbf{M}\mathbf{r}_2)' = \mathbf{r}_1'$. As a consequence, there exists a constant vector $\mathbf{x} = [x_0, y_0]$ such that $\mathbf{r}_1 = \mathbf{M}\mathbf{r}_2 + \mathbf{x}$.

□

This result might seem merely academic; but certain curves are *constructed in this way* for practical purposes. Roads leading to or from a highway frequently consist of parts with curvature *increasing* (or decreasing) *at a constant rate*. On such a road, the driver has to turn his/her driving wheel at a constant rate. An example of a function increasing with a constant rate is $\kappa(s) = s$. The resulting curve in Fig. 1.6 is called a *clothoid* curve – it has $\kappa(s) = s$ as the associated curvature function, and you have probably experienced it as a driver!

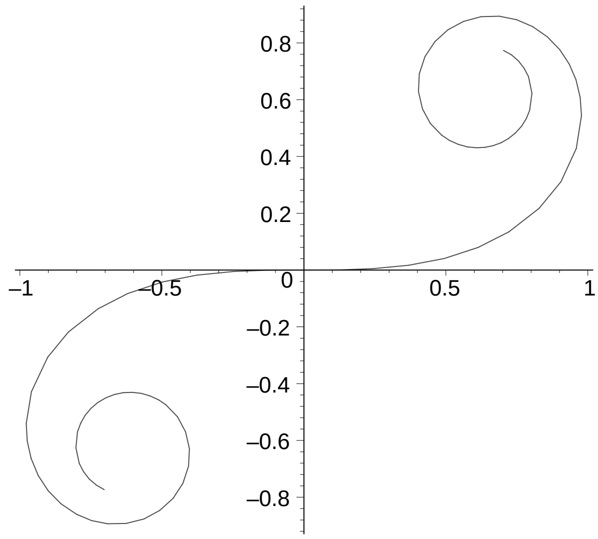


Figure 1.6: Curve with $\kappa(s) = s,\ -3 \leq s \leq 3$

Illustration 1.52 *Applet: Parametrization of κ, graph of κ, choice of coordinates of start point and start angle yield curve.*

1.3.6 The local canonical form of a plane curve

Locally, a smooth vector function can be approximated by correponding Taylor polynomials. Let us try to rediscover curvature of a plane curve in the 2nd degree Taylor polynomium of a (unit speed) parametrization $\mathbf{r} : I \to \mathbf{R}^2$. For simplicity, assume that $0 \in int(I)$, and (possibly after a rigid motion of the plane), that $\mathbf{r}(0) = \mathbf{0}$, that $\mathbf{r}'(0) = \mathbf{t}(0) = [1, 0]$, and that the curve has curvature κ at the origin. As a result of Def. 1.36, we obtain the 2nd derivative $\mathbf{r}''(0) = \kappa\widehat{\mathbf{t}}(0) = [0, \kappa]$. In particular, the 2nd degree Taylor polynomial is of the form

$$\mathbf{r}^{(2)}_{(0)}(s) = [0, 0] + [1, 0]s + [0, \kappa]\frac{s^2}{2} = [0, \frac{\kappa}{2}s^2].$$

This representation (plus the corresponding error term) is known as the *local canonical form* of the curve in a neighbourhood of $P_0 = O$. It represents a parabola with curvature κ at its top point at O. The sign of κ determines, whether this approximating parabola (and the original curve close to P_0) is situated in the upper or in the lower half-plane. This corresponds to the result in Cor. 1.40.

Illustration 1.53 *Curves and 2nd degree Taylor polynomials at various points.*

1.3.7 Exercises

1. Given three points $P, Q, R \in \mathbf{R}^2$, not contained in a line $l \subset \mathbf{R}^3$. Show:
   (a) There is a unique circle containing P, Q, R. Find the center $C(P, Q, R)$ and the radius $\rho(P, Q, R)$ of this circle.
   (b) Show that $C(P_{s_0}, P_{s_1}, P_{s_2})$ and $\rho(P_{s_0}, P_{s_1}, P_{s_2})$ converge when s_1 and s_2 tend to s_0.

2. (a) Let C be the (spiral) curve given by the regular parametrization $\mathbf{r} :]\varepsilon, \infty[\to \mathbf{R}^2$, $\mathbf{r}(t) = [t\cos t, t\sin t]$, $\varepsilon > 0$. Determine the velocity $\mathbf{r}'(t)$, the speed $v(t) = |\mathbf{r}'(t)|$ and the unit tangent vector field $\mathbf{t}(t)$ associated to this parametrization. Plot the x-coordinate of $\mathbf{t}(t)$. Determine an angle function $\theta :]\varepsilon, \infty[\to \mathbf{R}$ such that $\mathbf{t}(t) = [\cos(\theta(t)), \sin(\theta(t))]$. You cannot get the curvature function immediately by differentiating θ, since $\mathbf{r}$ is not a unit speed parametrization.
   (b) Calculate the associated curvature function $\kappa(t)$ and show that $\lim_{t\to\infty} \kappa(t) = 0$.
   (c) Determine a parametrization for Cs evolute curve E_C.

3. Let C denote a plane curve with *unit speed* parametrization and associated unit tangent and normal vector fields $\mathbf{t}(s)$ and $\mathbf{n}(s) = \hat{\mathbf{t}}(s)$. Show (using Prop. 1.10):
   (a) $\mathbf{n}'(s) \cdot \mathbf{n}(s) = 0$.
   (b) $\mathbf{n}'(s) \cdot \mathbf{t}(s) = -\kappa(s)$.
   (c) $\mathbf{n}'(s) = -\kappa(s)\mathbf{t}(s)$.

4. (a) Show Prop. 1.44 about evolute curves using Lemma 1.32.
   (b) Let C denote a regular plane curve with evolute E_C. Assume C given by a unit speed parametrization, and that κ is a smooth function of the parameter s. Show that the derivative of the associated parametrization for E_C at C_{P_s} is given by $-\frac{\kappa'(s)}{\kappa(s)}\mathbf{n}(s)$.

(c) Show that E_C is regular except at the points C_{P_s} corresponding to points $P_s \in C$ with critical values of κ, i.e., $\kappa'(s) = 0$.

(d) Let $P_s \in C$ and $C_{P_s} \in E_C$ be two associated points on C and its evolute curve E_C such that $\kappa'(s) \neq 0$. Show that the tangent lines to C at P and to E_C at $C_P$are perpendicular to each other.

(e) An exercise on involutes, e.g., the involute of a circle.

5. Reprove Prop. 1.48 using Thm. 1.51.

1.4 Space Curves

1.4.1 Moving frame, osculating plane and associated indicatrices

The moving frame

One of the major methods investigating curves (and later on surfaces) is to choose a nice perspective. In technical terms, we need *at every point* of the curve under investigation a basis of $\mathbf{R}^2$, resp. $\mathbf{R}^3$ adapted to the local situation. For a plane curve C given by a (unit speed) parametrization $\mathbf{r}$ and a point $P \in C$ with $\overrightarrow{OP} = \mathbf{r}(s)$, this *moving frame* is given by the unit tangent vector $\mathbf{t}(s)$ and its hat vector $\widehat{\mathbf{t}}(s)$, providing us in fact with a moving *orthonormal* basis of $\mathbf{R}^2$ at every point of the curve.

For a space curve C given by a (unit speed) parametrization $\mathbf{r}$, we need *three* (mutually orthogonal unit) basis vectors at every point $P \in C$. One of them should again be the tangent vector to the curve. What to choose for the two other basis vectors? In the following, we have to assume, that the point $P \in C$ (with $\overrightarrow{OP} = \mathbf{r}(s)$) is a non-inflection point; cf. Def. 1.35. In that case, we are provided with the *principal normal vector* $\mathbf{n}(s)$ given by $\mathbf{t}'(s) = \kappa(s)\mathbf{n}(s)$. By definition, the principal normal vector $\mathbf{n}(s)$ is a unit vector; moreover, it is perpendicular to the unit tangent vector $\mathbf{t}(s)$ at P by Rem. 1.37. This leaves us with almost no choice for the third vector in an orthonormal basis:

Definition 1.54 *Let $\mathbf{r} : I \to \mathbf{R}^3$ denote a unit speed parametrization of a space curve C; assume that $P_s \in C$ given by $\overrightarrow{OP_s} = \mathbf{r}(s)$ is a non-inflection point. The moving frame at P_s consists of*

- *the* unit tangent vector $\mathbf{t}(s) = \mathbf{r}'(s)$*;*

- *the* principal normal vector $\mathbf{n}(s) = \frac{\mathbf{r}''(s)}{|\mathbf{r}''(s)|}$;
- *the* binormal vector $\mathbf{b}(s) = \mathbf{t}(s) \times \mathbf{n}(s)$.

Illustration 1.55 *Moving frame moving along a space curve choosen by the user.*

For explicit calculations using a regular (non unit-speed) parametrization of a space curve C, we have to find formulas for the three vectors in the moving frame: Let $\mathbf{r}$ denote such a regular parametrization, and let $\mathbf{t}(t), \mathbf{n}(t), \mathbf{b}(t)$ denote the moving frame vectors at a non-inflection point $P_t \in C$ with $\overrightarrow{OP_t} = \mathbf{r}(t)$.

Remark 1.56 *1. The planes* $sp(\mathbf{r}'(t), \mathbf{r}''(t))$ *and* $sp(\mathbf{t}(t), \mathbf{n}(t))$ *agree; this is an easy consequence of (1.7).*

2. The determination of the moving frame at the point P_t *with* $\overrightarrow{OP_t} = \mathbf{r}(t)$ *is thus usually performed in the following three steps:*

$$1.\ \mathbf{t}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|};$$
$$2.\ \mathbf{b}(t) = \frac{\mathbf{r}'(t) \times \mathbf{r}''(t)}{|\mathbf{r}'(t) \times \mathbf{r}''(t)|};$$
$$3.\ \mathbf{n}(t) = \mathbf{b}(t) \times \mathbf{t}(t).$$

Example 1.57 *Let us again look at a helix with parametrization* $\mathbf{r}(t) = [a\cos t, a\sin t, bt]$, $a, b > 0$. *(cf. Ex. 1.4 and 1.43). Using the formulas from Rem. 1.56, we obtain at* P_t *with* $\overrightarrow{OP_t} = \mathbf{r}(t)$:

$$\begin{aligned} \mathbf{t}(t) &= \frac{1}{\sqrt{a^2+b^2}}[-a\sin t, a\cos t, b]; \\ \mathbf{b}(t) &= \frac{1}{\sqrt{a^2+b^2}}[b\sin t, -b\cos t, a]; \\ \mathbf{n}(t) &= [-\cos t, -\sin t, 0]. \end{aligned} \tag{1.17}$$

In particular, at every point P_t *along the curve, the principal normal vector* $\mathbf{n}(t)$ *is parallel to the* XY*-plane pointing from the curve in direction of the* $z-$*axis.*

Similar to Sect. 1.3.3, one can visualise the moving frame by three *curves*, their *indicatrices*. This time, those curves will be contained in the unit 2-sphere $S^2 := \{[x, y, z] \in \mathbf{R}^3 | x^2 + y^2 + z^2 = 1\} \subset \mathbf{R}^3$; the tangent, principal and binormal vector fields can be viewed as maps $\mathbf{t}, \mathbf{n}, \mathbf{b} : I \to S^2$; they have moreover the property of hitting mutually perpendicular vectors at any time.

Illustration 1.58 *3 indicatrices of the same curve (chosen by the user) marked by points in 3 colours travelling on a unit sphere while another point moves along the curve.*

Normal and osculating planes

Each pair of two out of the three vectors in the moving frame of a space curve C span a plane; at least two of them deserve to be mentioned here. We get a more adequate geometric picture when these planes are translated into parallel planes through the point $P \in C$ at which they are defined (just as the tangent line at P is obtained from the 1-dimensional subspace $sp(\mathbf{t})$ by translation into P).

Let again $\mathbf{r} : I \to \mathbf{R}^3$ denote a regular parametrization of a space curve C and $P_t \in C$ denote the point P_t with $\overrightarrow{OP_t} = \mathbf{r}(t)$. The *normal plane* η_t at P_t is the plane through P_t normal to the tangent vector $\mathbf{t}(t)$, or equivalently, with parallel plane $sp(\mathbf{n}(t), \mathbf{b}(t))$.

The *osculating plane* ω_t at P_t is the plane through the non-inflection point P_t normal to the binormal vector $\mathbf{b}(t)$, or equivalently, with parallel plane $sp(\mathbf{n}(t), \mathbf{b}(t))$, or equivalently, with parallel plane $sp(\mathbf{r}'(t), \mathbf{r}''(t))$; cf. Rem. 1.56.1. It is in fact the *best approximating plane* to the curve near P_t.

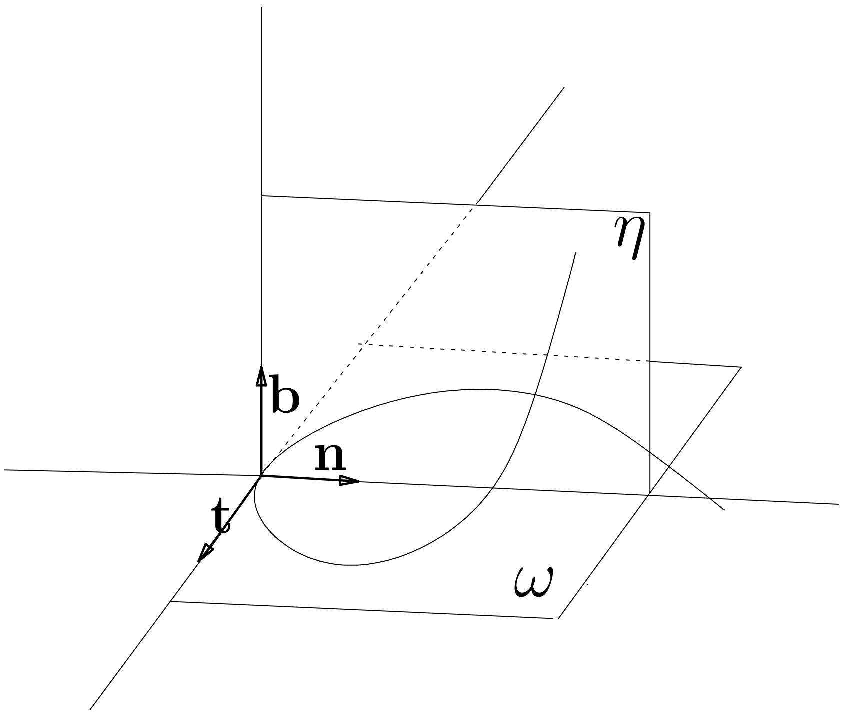


Figure 1.7: Moving frame, osculating plane and normal plane

1. It contains both the tangent line, and *for every* regular parametrization the line through P_t parallel to the acceleration vector $\mathbf{r}''(t)$. In a physical sense,

the kinematic forces pull within the osculation plane regardless the speed with which the curve is run through.

2. An alternative and geometric appealing construction of the osculating plane ω_t at P_t is as follows: Choose two points $P_{t_1}, P_{t_2} \in C$ close to P_t. If the three points P_t, P_{t_1}, P_{t_2} are not contained in a line, they determine a plane $\omega(t, t_1, t_2) \subset \mathbf{R}^3$ containing those three points. When t_1 and t_2 tend to t, these plane tends to a limit plane, which is the osculating plane; cf. Exc. 1.4.5.?

Illustration 1.59 *Moving normal and osculating planes along a (user chosen) space curve.*

Proposition 1.60 *A point* $\mathbf{x} = [x, y, z]$ *is contained in the osculating plane* ω_t *at* $P_t \in C$ *with* $\overrightarrow{OP_t} = \mathbf{r}(t)$ *if and only if*

$$\mathbf{b}(t) \cdot \mathbf{x} = \mathbf{b}(t) \cdot \mathbf{r}(t),$$

or equivalently, if the matrix $\begin{bmatrix} \mathbf{x} - \mathbf{r}(t) \\ \mathbf{r}'(t) \\ \mathbf{r}''(t) \end{bmatrix}$ *is singular.*

Proof: Exc. 1.4.5.?.

□

Example 1.61 *The osculating plane to the helix from Ex. 1.57 at* P_t *is given by the equation*

$$\mathbf{b}(t) \cdot [x, y, z] = \mathbf{b}(t) \cdot \overrightarrow{OP_0}, \text{ i.e.}$$

$$b \sin t \cdot x - b \cos t \cdot y + a \cdot z = abt$$

or by the parametrization

$$\mathbf{r}(u, v) = [a \cos t, a \sin t, bt] + u[-a \sin t, a \cos t, b] + v[-a \cos t, -asint, 0], \; (u, v) \in \mathbf{R}^2.$$

Also for a space curve, the *osculating circle* along a space curve C at $P \in C$ is the circle best approximating C at P. The circle through three points on C is of course contained in the plane determined by the three points. As the limit circle, the osculating circle at P is contained in the limit position for those planes, i.e., the osculating plane. Given a parametrization $\mathbf{r}$ for the curve, we have thus the following analogue of Prop. 1.44 determining the osculating circle at P:

Proposition 1.62 *The osculating circle at P_t with $\overrightarrow{OP_t} = \mathbf{r}(t)$ is contained in the osculating plane ω_t determined by Prop. 1.60. It has radius $\rho(t) = \frac{1}{\kappa(t)}$, center C_t with $\overrightarrow{P_tC_t} = \frac{1}{\kappa(t)}\mathbf{n}(t)$.*

Illustration 1.63 *Osculating circle along a space curve (like the one done by Robert).*

1.4.2 Torsion and Frenet's equations

Let $\mathbf{r} : I \to \mathbf{R}^3$ denote the unit-speed parametrization of a curve, and let $\mathbf{t}, \mathbf{n}, \mathbf{b} : I \to \mathbf{R}^3$ the associated moving frame (unit tangent, principal normal and binormal vector fields). At every (non-inflection) point $\mathbf{r}(s)$ along the curve, we can form the 3×3 Frenet-matrix

$$\mathbf{F}(s) = [\mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s)]$$

with the three moving frame vectors as the *column* vectors of $\mathbf{F}(s)$. Since the three vectors are mutually perpendiuclar unit vectors, the matrix $\mathbf{F}(s)$ is an *orthogonal* matrix for every $s \in I$, i.e., it satisfies

$$\mathbf{I} = \mathbf{F}^T(s)\mathbf{F}(s) = \mathbf{F}(s)\mathbf{F}^T(s). \tag{1.18}$$

Moreover, $\mathbf{F}(s)$ has determinant 1, since $\det \mathbf{F}(s) = (\mathbf{t}(s) \times \mathbf{n}(s)) \cdot \mathbf{b}(s) = \mathbf{b}(s) \cdot \mathbf{b}(s) = 1$. A matrix satisfying (1.18) with determinant 1 is called *special orthogonal*; we write $\mathbf{F}(s) \in SO(3)$.

To see, how the moving frame changes along the curve given by the parametrization $\mathbf{r}$, we want to determine the derivative of the matrix $\mathbf{F}(s)$, i.e., the matrix

$$\mathbf{F}'(s) = [\mathbf{t}'(s), \mathbf{n}'(s), \mathbf{b}'(s)]$$

with the differentials of the moving frame vectors as columns.

Since $\mathbf{F}(s)$ is orthogonal, it is also invertible; in fact $\mathbf{F}^{-1}(s) = \mathbf{F}^T(s)$. Hence, with the matrix $\mathbf{A}(s) = \mathbf{F}^T(s)\mathbf{F}'(s)$, we get an equation

$$\mathbf{F}'(s) = \mathbf{F}(s)\mathbf{F}^T(s)\mathbf{F}'(s) = \mathbf{F}(s)\mathbf{A}(s). \tag{1.19}$$

We want to determine the matrix $\mathbf{A}(s)$ relating $\mathbf{F}(s)$ and $\mathbf{F}'(s)$:

Lemma 1.64 *The matrix $\mathbf{A}(s)$ is* skew-symmetric *for every $s \in I$, i.e.,*

$$\mathbf{A}(s) + \mathbf{A}^T(s) = \mathbf{0}.$$

Proof: By definition, $\mathbf{A}(s) = \mathbf{F}^T(s)\mathbf{F}'(s)$ and hence $\mathbf{A}^T(s) = \mathbf{F}'^T(s)\mathbf{F}(s)$. As a result,

$$\mathbf{A}(s) + \mathbf{A}^T(s) = \mathbf{F}^T(s)\mathbf{F}'(s) + \mathbf{F}'^T(s)\mathbf{F}(s) = \frac{d}{ds}(\mathbf{F}^T(s)\mathbf{F}(s)) = \frac{d}{ds}\mathbf{I} = \mathbf{0}.$$

□

It is easy to determine the 1st column of the matrix $A(s)$: By the definition of curvature, we have (Def. 1.36): $\mathbf{t}'(s) = 0{\cdot}\mathbf{t}(s) + \kappa(s){\cdot}\mathbf{n}(s) + 0{\cdot}\mathbf{b}(s)$. Combining this with the skew-commutativity of the matrix $A(s)$ (Lemma 1.64), we can determine most of the entries of that matrix, to wit:

$$A(s) = \begin{bmatrix} 0 & -\kappa(s) & 0 \\ \kappa(s) & 0 & -?(s) \\ 0 & ?(s) & 0 \end{bmatrix}.$$

Definition 1.65 *Let $P_s \in C$ denote a non-inflection point on a regular space curve C given by a unit speed parametrization* $\mathbf{r}$ *such that $\overrightarrow{OP_s} = \mathbf{r}(s)$. The* torsion *of C at P_s is defined by the equation*

$$\mathbf{b}'(s) = -\tau(s)\mathbf{n}(s). \tag{1.20}$$

Before interpreting this new invariant, let us note that we have proved

Theorem 1.66 *(Frenet's equations)*

$$\begin{aligned} \mathbf{t}'(s) &= \qquad\qquad \kappa(s)\mathbf{n}(s) \\ \mathbf{n}'(s) &= -\kappa(s)\mathbf{t}(s) \qquad\qquad + \tau(s)\mathbf{b}(s) \\ \mathbf{b}'(s) &= \qquad\qquad -\tau(s)\mathbf{n}(s) \end{aligned}.$$

Corollary 1.67

$$A(s) = \begin{bmatrix} 0 & -\kappa(s) & 0 \\ \kappa(s) & 0 & -\tau(s) \\ 0 & \tau(s) & 0 \end{bmatrix}.$$

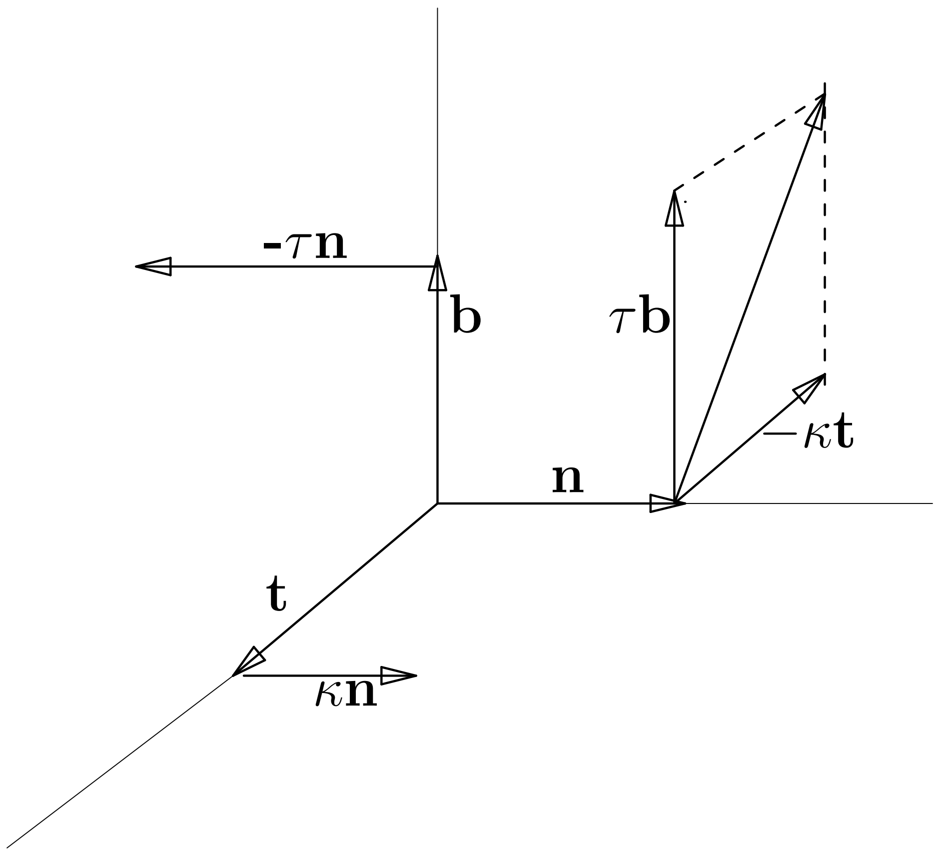


Figure 1.8: Frenet's equations

Torsion: Interpretation and Calculation

The binormal vector $\mathbf{b}(s)$ is by definition a unit vector orthogonal to the osculating plane ω_{P_s}. Hence, the torsion as (the negative of) the rate of change of $\mathbf{b}(s)$ is equivalent to the rate of change of ω_{P_s}; hence it measures, how quickly the osculating plane changes when moving away from P_s along the curve, or equivalently, how quickly the curve disappears or twists away from ω_{P_s}, in particular:

Proposition 1.68 *A space curve C without inflection points is contained in a plane if and only if its torsion is 0 everywhere.*

Proof: Let $\mathbf{r}: I \to \mathbf{R}^3$ denote a unit speed parametrization of the curve C.

If C is contained in a plane $\pi \subset \mathbf{R}^3$, then $\mathbf{r}(s) - \mathbf{r}(0)$ and all derivatives $\mathbf{r}^i(s)$ are contained in a parallel 2-dimensional vector space $\alpha \subset \mathbf{R}^3$. In particular, the binormal $\mathbf{b}(s)$ is a continuous vector function with values among the two vectors of length one perpendicular to α. Hence, $\mathbf{b}(s)$ is a constant vector, $\mathbf{b}'(s) = \mathbf{0}$, and hence $\tau(s) \equiv 0$.

From $\tau(s) = 0$, we conclude from the last equation in Thm. 1.66, that $\mathbf{b}'(s) = \mathbf{0}$, and hence that the binormal vector $\mathbf{b}(s) = \mathbf{b}$ is constant along C. Let $s_0 \in I$. The plane through $\mathbf{r}(s_0)$ perpendicular to $\mathbf{b}$ is given by the equation

$$(\mathbf{x} - \mathbf{r}(s_0)) \cdot \mathbf{b} = 0. \qquad (1.21)$$

We show, that all vectors $\mathbf{x} = \mathbf{r}(s),\ x \in I$ satisfy (1.21): This is obviously true for $\mathbf{x} = \mathbf{r}(s_0)$. Moreover, the function $(\mathbf{r}(s) - \mathbf{r}(s_0)) \cdot \mathbf{b}$ has derivative $\mathbf{r}'(s) \cdot \mathbf{b} = \mathbf{t}(s) \cdot \mathbf{b} = \mathbf{0}$, i.e., it is constant with constant 0.

□

Remark 1.69 *The curvature of a space curve C at a point P was defined to be non-negative; its torsion can be both* positive, negative or zero. *The sign in (1.20) is somehow arbitrary, and some authors write $\tau(s)$ instead of our $-\tau(s)$.*

Frenet's equations (Thm. 1.66) can be used to derive a formula for the torsion $\tau(P)$ at a point P on our curve C in terms of the first three derivatives of a parametrization $\mathbf{r}$ of the curve.

Proposition 1.70 *Let C be a curve with regular parametrization $\mathbf{r}$, and let $P_t \in C$ be a non-inflection point given by $\overrightarrow{OP} = \mathbf{r}(t)$. Then,*

$$\tau(P) = \frac{(\mathbf{r}'(t) \times \mathbf{r}''(t)) \cdot \mathbf{r}'''(t)}{|\mathbf{r}'(t) \times \mathbf{r}''(t)|^2} = \frac{[\mathbf{r}'(t), \mathbf{r}''(t), \mathbf{r}'''(t)]}{|\mathbf{r}'(t) \times \mathbf{r}''(t)|^2}.$$

(The last numerator is the determinant of the matrix with the three derivatives as row vectors.)

Proof: We reuse the formulas (1.7) for the velocity and the acceleration vector in the following form:

$$\begin{aligned}\mathbf{r}'(t) &= v(t)\mathbf{t}(s(t));\\ \mathbf{r}''(t) &= v'(t)\mathbf{t}(s(t)) + v^2(t)\kappa(s(t))\mathbf{n}(s(t)).\end{aligned} \tag{1.22}$$

A calculation of their cross product yields:

$$\mathbf{r}'(t) \times \mathbf{r}''(t) = v^3(t)\kappa(s(t))\mathbf{b}(s(t)), \tag{1.23}$$

and thus,

$$(\mathbf{r}'(t) \times \mathbf{r}''(t)) \cdot \mathbf{r}'''(t) = v^3(t)\kappa(s(t))(\mathbf{b}(s(t)) \cdot \mathbf{r}'''(t)). \tag{1.24}$$

To calculate $\mathbf{b}(s(t)) \cdot \mathbf{r}'''(t)$, note first that $\mathbf{b}(s(t)) \cdot \mathbf{r}''(t) = 0$, since $\mathbf{r}''(t)$ is contained in the osculating plane ω_P at P, cf. Rem. 1.56. Hence, we may use first the fundamental trick (Lemma 1.10) and then the last of Frenet's equations, cf. Thm. 1.66, to obtain:

$$\mathbf{b}(s(t)) \cdot \mathbf{r}'''(t) = -v(t)\mathbf{b}'(s(t)) \cdot \mathbf{r}''(t) = v(t)\tau(s(t))(\mathbf{n}(s(t)) \cdot \mathbf{r}''(t)).$$

Applying the second equation in (1.22) once again, we derive:

$$\mathbf{b}(s(t)) \cdot \mathbf{r}'''(t) = v^3(t)\kappa(s(t))\tau(s(t)). \tag{1.25}$$

Substituting (1.25) into (1.24), we obtain:

$$(\mathbf{r}'(t) \times \mathbf{r}''(t)) \cdot \mathbf{r}'''(t)) = v^3(t)\kappa(s(t))(\mathbf{b}(s(t)) \cdot \mathbf{r}'''(t)) = v^6(t)\kappa^2(s(t))\tau(s(t)).$$

On the other hand, (1.23) tells us, that $|\mathbf{r}'(t) \times \mathbf{r}''(t)|^2 = v^6(t)\kappa^2(s(t))$, which implies the formula in Prop. 1.70 for the torsion $\tau(P_t) = \tau(s(t))$.

□

Remark 1.71 *From (1.25), we deduct that the torsion $\tau(P_t)$ has the same sign as the entity $\mathbf{b}(P)\cdot\mathbf{r}'''(s)$ – since both speed v and curvature κ are positive entities. this observation can be given the following interpretation:*

Euclidean space $\mathbf{E}^3$ is divided into two half-spaces *by the osculating plane ω_P. The torsion τ_P is positive, if and only if $\mathbf{r}'''(t)$ lies in the half space that $\mathbf{b}(P)$ points into, i.e., if the piece of curve given by $\mathbf{r}(t+\varepsilon)$ for small values $\varepsilon > 0$ is contained in that half-space. It is negative, if and only if $\mathbf{r}'''(t)$ and thus the piece of curve given by $\mathbf{r}(t+\varepsilon)$ for small values $\varepsilon > 0$ is contained on the opposit half-space. The absolute value of the torsion $\tau(P)$ measures, given $\kappa(P)$,* how fast *the curve twists away from ω_P into one or the other half-space. See also Sect. 1.4.4.*

Example 1.72 *For the helix from Ex. 1.4 with parametrization* $\mathbf{r} : \mathbf{R} \to \mathbf{R}^3$, $\mathbf{r}(t) = [a\cos t, a\sin t, bt]$, *we calculate using the results of Ex. 1.43.3:*

$$\begin{aligned} \mathbf{r}'''(t) &= [a\sin t, -a\cos t, 0]; \\ \mathbf{r}'(t) \times \mathbf{r}''(t) &= [ab\sin t, -ab\cos t, a^2]; \\ (\mathbf{r}'(t) \times \mathbf{r}''(t)) \cdot \mathbf{r}'''(t) &= a^2 b; \\ \tau(P) &= \frac{b}{a^2 + b^2}. \end{aligned} \tag{1.26}$$

Note, that both curvature (cf. (1.16)) and torsion are constant *along the helix.*

Illustration 1.73 *Space curve with attached curvature and torsion plots and attached tangent and binormal indicatrix.*

1.4.3 Curvature and torsion functions characterize a space curve

We saw in Sect. 1.3.5 that a *plane* curve is characterized (up to a rigid motion) by its curvature function. This is no longer true for a space curve: A circle and a helix can have the same constant curvature function, cf.Ex. 1.43.3. But a result similar to Thm. 1.51holds if the torsion function is adjoined as an additional invariant:

Theorem 1.74 ***(The fundamental theorem for space curves)*** *Let* $\bar{\kappa}, \bar{\tau} : (a, b) \to \mathbf{R}$ *denote smooth functions with* $\bar{\kappa}(s) > 0$ *for all* $a < s < b$. *Then, there is a curve with unit speed parametrization* $\mathbf{r} : (a, b) \to \mathbf{R}^3$, *whose curvature and torsion functions are* $\bar{\kappa}$, *resp.* $\bar{\tau}$. *Furthermore, any two such curves differ by a proper rigid motion.*

Proof:

Existence Consider the function

$$\mathbf{A}(s) = \begin{bmatrix} 0 & -\bar{\kappa}(s) & 0 \\ \bar{\kappa}(s) & 0 & -\bar{\tau}(s) \\ 0 & \bar{\tau}(s) & 0 \end{bmatrix}$$

with $\bar{\kappa}(s)$ and $\bar{\tau}(s)$ as in Thm. 1.74. Remark that $\mathbf{A}(s)$ is a *skew-symmetric matrix* for every $a < s < b$.

In our first step, we attempt to find a moving frame $\{\bar{\mathbf{t}}(s), \bar{\mathbf{n}}(s), \bar{\mathbf{b}}(s)\}$ satisfying the Frenet equations from Thm. 1.66. With

$$\mathbf{F}(s) = [\bar{\mathbf{t}}(s), \bar{\mathbf{n}}(s), \bar{\mathbf{b}}(s)]$$

we look for a solution of the system of *linear* differential equations that in matrix form reads as:

$$\mathbf{F}'(s) = \mathbf{F}(s)\mathbf{A}(s). \tag{1.27}$$

That system consists of one *linear!* differential equation for each entry in the matrix $\mathbf{F}'(s)$, nine in total. By an important result in the theory of ordinary differential equations, such a system has always a unique solution $\{\bar{\mathbf{t}}(s), \bar{\mathbf{n}}(s), \bar{\mathbf{b}}(s)\}$ defined on the *entire interval* (a, b) with given initial values $[\bar{\mathbf{t}}(s_0), \bar{\mathbf{n}}(s_0), \bar{\mathbf{b}}(s_0)]$.

What initial values $[\bar{\mathbf{t}}(s_0), \bar{\mathbf{n}}(s_0), \bar{\mathbf{b}}(s_0)]$ may one choose "legally"? Well, they ought better arise as a moving frame, i.e., the matrix

$$\mathbf{F}(s_0) = [\bar{\mathbf{t}}(s_0), \bar{\mathbf{n}}(s_0), \bar{\mathbf{b}}(s_0)]$$

needs to be a *special orthogonal* matrix $\mathbf{F}(s_0) \in SO(3)$, i.e., $\mathbf{F}(s_0)^T\mathbf{F}(s_0) = \mathbf{I}$ and $\det \mathbf{F}(s_0) = 1$.

Lemma 1.75 *The matrix solutions* $\mathbf{F}(s)$ *of (1.27) are special orthogonal* for all s, *i.e.,* $\mathbf{F}(s)\mathbf{F}^T(s) = \mathbf{I}$ *and* $\det \mathbf{F}(s) = 1$ *for all* $a < s < b$.

Proof: (of Lemma 1.75:) We want to show that the matrix-valued function $\mathbf{G}(s) = \mathbf{F}(s)\mathbf{F}^T(s)$ takes the constant value $\mathbf{G}(s) = \mathbf{I}$ for all $a < s < b$. To obtain this result, we differentiate $\mathbf{G}(s)$ and obtain:

$$\mathbf{G}'(s) = \mathbf{F}'(s)\mathbf{F}^T(s) + \mathbf{F}(s)\mathbf{F}'^T(s).$$

Since $\mathbf{F}(s)$ solves (1.27) above, we have

$$\mathbf{G}'(s) = \mathbf{F}(s)\mathbf{A}(s)\mathbf{F}^T(s) + \mathbf{F}(s)\mathbf{A}^T(s)\mathbf{F}^T(s) = \mathbf{F}(s)(\mathbf{A}(s) + \mathbf{A}^T(s))\mathbf{F}^T(s) = \mathbf{0}.$$

(The last equation relies on $\mathbf{A}(s)$ being skew-symmetric.) We conclude that the matrix-valued function $\mathbf{G}(s)$ is constant and takes the value $\mathbf{G}(s_0) = \mathbf{I}$ everywhere on the interval (a, b). The determinant of a matrix function is continuous. Since the determinant of an orthogonal matrix can only take the values 1 or -1, it must take the constant value $\det \mathbf{F}(s) = \det \mathbf{F}(s_0) = 1$.

□

Using the first column $\bar{\mathbf{t}}(s)$ of the solution matrix $F(s)$, we define the (unit-speed parametrization! of the) curve $\mathbf{r}(s)$, by integration:

$$\mathbf{r}(s) = \int_{s_0}^{s} \bar{\mathbf{t}}(t)dt. \tag{1.28}$$

Let us find first the moving frame $\{\mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s)\}$ and then the curvature and torsion functions $\kappa(s)$, resp. $\tau(s)$ associated to the curve given by 6.3:

Lemma 1.76 $\mathbf{t}(s) = \bar{\mathbf{t}}(s)$, $\mathbf{n}(s) = \bar{\mathbf{n}}(s)$, $\mathbf{b}(s) = \bar{\mathbf{b}}(s)$.
$\kappa(s) = \bar{\kappa}(s), \tau(s) = \bar{\tau}(s)$ for all $a < s < b$.

Proof: (of Lemma 1.76:) By the fundamental theorem of Analysis, $\mathbf{t}(s) = \mathbf{r}'(s) = \bar{\mathbf{t}}(s)$. In particular, $\mathbf{r}$ is a unit speed parametrization. Moreover, $\kappa(s)\mathbf{n}(s) = \mathbf{t}'(s) = \bar{\mathbf{t}}'(s) = \bar{\kappa}(s)\bar{\mathbf{n}}(s)$. The first equation reflects the definition of curvature, the last one the fact that $\bar{\mathbf{t}}(s)$ is the first column in the solution of (1.27). The vectors $\mathbf{n}(s)$ and $\bar{\mathbf{n}}(s)$ are both unit vectors – by definition, resp. by Lemma 1.75. Since $\kappa(s)$ and $\bar{\kappa}(s)$ are both positive for all s, we conclude that $\mathbf{n}(s) = \bar{\mathbf{n}}(s)$ and hence $\kappa(s) = \bar{\kappa}(s)$.

Since $\mathbf{F}(s)$ is *special orthogonal*, $\mathbf{b}(s) = \mathbf{t}(s) \times \mathbf{n}(s) = \bar{\mathbf{t}}(s) \times \bar{\mathbf{n}}(s) = \bar{\mathbf{b}}(s)$. From the definition of torsion, resp. from (1.27), we conclude:
$-\tau(s)\mathbf{n}(s) = \mathbf{b}'(s) = \bar{\mathbf{b}}'(s) = -\bar{\tau}(s)\bar{\mathbf{n}}(s)$ and hence: $\tau(s) = \bar{\tau}(s)$.

□

Uniqueness Let $\mathbf{r}_1, \mathbf{r}_2 : (a, b) \to \mathbf{R}^3$ denote two unit-speed parametrizations with the same associated curvature and torsion functions, i.e., $\kappa_1(s) = \kappa_2(s)$ and $\tau_1(s) = \tau_2(s)$ for all $s \in (a, b)$. The associated Frenet matrices

$$\mathbf{F}_1(s) = [\mathbf{t}_1(s), \mathbf{n}_1(s), \mathbf{b}_1(s)] \text{ and } \mathbf{F}_2(s) = [\mathbf{t}_2(s), \mathbf{n}_2(s), \mathbf{b}_2(s)]$$

are both contained in $SO(3)$, and they satisfy equations (1.19)

$$\mathbf{F}_1'(s) = \mathbf{F}_1(s)\mathbf{A}(s) \text{ and } \mathbf{F}_2'(s) = \mathbf{F}_2(s)\mathbf{A}(s)$$

with the *same* matrix function

$$A(s) = \begin{bmatrix} 0 & -\kappa(s) & 0 \\ \kappa(s) & 0 & -\tau(s) \\ 0 & \tau(s) & 0 \end{bmatrix}.$$

Consider the matrix function $\mathbf{M}(s) = \mathbf{F}_2(s)\mathbf{F}_1^{-1}(s) = \mathbf{F}_2(s)\mathbf{F}_1^T(s)$. We wish to show that $\mathbf{M}(s)$ is *constant* and differentiate:

$$\begin{aligned}\mathbf{M}'(s) &= \mathbf{F}_2'(s)\mathbf{F}_1^T(s) + \mathbf{F}_2(s)\mathbf{F}_1'^T(s) = \mathbf{F}_2(s)\mathbf{A}(s)\mathbf{F}_1^T(s) + \mathbf{F}_2(s)\mathbf{A}^T(s)\mathbf{F}_1^T(s) \\ &= \mathbf{F}_2(s)(\mathbf{A}(s) + \mathbf{A}^T(s))\mathbf{F}_1^T(s) = \mathbf{0}.\end{aligned}$$

Hence $\mathbf{M}(s) = \mathbf{M} \in SO(3)$ is a constant special orthogonal matrix satisfying $\mathbf{M}\mathbf{F}_1(s) = \mathbf{F}_2(s)\mathbf{F}_1^T(s)\mathbf{F}_1(s) = \mathbf{F}_2(s)$, and in particular

$$\mathbf{M}\mathbf{t}_1(s) = \mathbf{t}_2(s), \mathbf{M}\mathbf{n}_1(s) = \mathbf{n}_2(s), \mathbf{M}\mathbf{b}_1(s) = \mathbf{b}_2(s).$$

The first of these equations shows that

$$(\mathbf{M}\mathbf{r}_1(s))' = \mathbf{M}(\mathbf{r}_1(s))' = \mathbf{M}\mathbf{t}_1(s) = \mathbf{t}_2(s) = \mathbf{r}_2(s)'.$$

In particular, the vector function $s \mapsto \mathbf{r}_2(s) - \mathbf{M}\mathbf{r}_1(s)$ has derivative $\mathbf{0}$, and hence the vector $\mathbf{v} = \mathbf{r}_2(s) - \mathbf{M}\mathbf{r}_1(s)$ is *constant.* This leads to:

$$\mathbf{r}_2(s) = \mathbf{M}\mathbf{r}_1(s) + \mathbf{v},$$

i.e. the curve given by $\mathbf{r}_2(s)$ is the result of a rotation (given by the matrix $\mathbf{M}$) and a translation (given by the vector $\mathbf{v}$) performed on the curve given by $\mathbf{r}_1(s)$.

□

1.4.4 The local canonical form of a space curve

Here is yet another way to grasp the meaning of the curvature and the torsion of a (space) curve C. For simplicity, we assume C given a unit speed parameterisation $\mathbf{r} : I \to \mathbf{R}^3$ with I an interval containing $s = 0$ in its interior. Furthermore, we assume (after a rigid motion in space), that $P_0 = O$ and that the Frenet moving frame vectors at O coincides with the coordinate unit vectors, i.e., $\mathbf{t}(0) = \mathbf{i}$, $\mathbf{n}(0) = \mathbf{j}$, and $\mathbf{b}(0) = \mathbf{k}$. We want to analyse the curve C close to the point $P_0 \in \mathbf{E}^3$ by looking at the 3rd degree Taylor polynomium of the (components of the) vector function $\mathbf{r}$ at O:

$$\mathbf{r}_0^{(3)}(s) = \mathbf{r}(0) + s\mathbf{r}'(0) + \frac{s^2}{2}\mathbf{r}''(0) + \frac{s^3}{6}\mathbf{r}'''(0). \tag{1.29}$$

By our assumptions and using Frenet's 2nd equation from Thm. 1.66 we obtain[6]:

$$\begin{aligned} \mathbf{r}(0) &= \mathbf{0}, \\ \mathbf{r}'(0) = \mathbf{t}(0) &= [1,0,0], \\ \mathbf{r}''(0) = \mathbf{t}'(0) = \kappa\mathbf{n}(0) &= [0,\kappa,0], \\ \mathbf{r}'''(0) = (\kappa\mathbf{n})'(0) = \kappa'\mathbf{n}(0) + \kappa\mathbf{n}'(0) = \kappa'\mathbf{n}(0) + \kappa(-\kappa\mathbf{t}(0) + \tau\mathbf{b}(0)) &= [-\kappa^2, \kappa', \kappa\tau]. \end{aligned}$$

Substituting this into (1.29), we get for the 3rd degree Taylor polynomial:

$$\mathbf{r}_0^{(3)}(s) = [s - \frac{\kappa^2 s^3}{6}, \frac{\kappa s^2}{2} + \frac{\kappa' s^3}{6}, \frac{\kappa\tau s^3}{6}].$$

This representation (plus the corresponding error term) is known as the *local canonical form* of the curve in a neighbourhood of $P_0 = O$. Remark, that the torsion only enters as a coordinate of $\mathbf{b}$. Since κ is assumed to be positive (compare Rem. 1.71):

Corollary 1.77 *The osculating plane ω_O divides Euclidean space into two half-spaces. If $\tau > 0$, the curve C runs into the half-space containing the binormal vector $\mathbf{b}$ for $t > 0$; if $\tau < 0$, it leaves that half-space for $t > 0$.*

The following figure shows a curve with given values κ and τ and its projections on the three planes spanned by two out of three of the vectors $\{\mathbf{t}, \mathbf{n}, \mathbf{b}\}$ from the Frenet frame:

1.4.5 Exercises

1. Given three points $P, Q, R \in \mathbf{R}^3$, not contained in a line $l \subset \mathbf{R}^3$. Show:
   (a) There is a unique plane $\omega_{P,Q,R}$ containing P, Q, R. Determine a unit normal vector $\mathbf{N}_{P,Q,R}$ to $\omega_{P,Q,R}$.
   (b) Let $P = P_{s_0}, Q = P_{s_1}$ and $R = P_{s_2}$ as in Sect. 1.3.1and Sect. 1.4.1. Show that the planes $\omega_{P_{s_0},P_{s_1},P_{s_2}}$ converge to a limit plane ω_{s_0} (the osculating plane).
   (Hint: Consider the planes $sp(\mathbf{r}(s_1) - \mathbf{r}(s_0), \mathbf{r}(s_2) - \mathbf{r}(s_1))$ and $sp(\frac{\mathbf{r}(s_1)-\mathbf{r}(s_0)}{s_1-s_0}, \frac{\frac{\mathbf{r}(s_2)-\mathbf{r}(s_1}{s_2-s_1} - \frac{\mathbf{r}(s_1)-\mathbf{r}(s_0)}{s_1-s_0}}{s_1-s_0})$. Show that the two planes agreee, and that the latter converges to $sp(\mathbf{r}'(s_0), \mathbf{r}''(s_0))$.
2. Show Prop. 1.60.

[6]We use the abbreviations: $\kappa = \kappa(0), \kappa' = \kappa'(0)$, and $\tau = \tau(0)$.

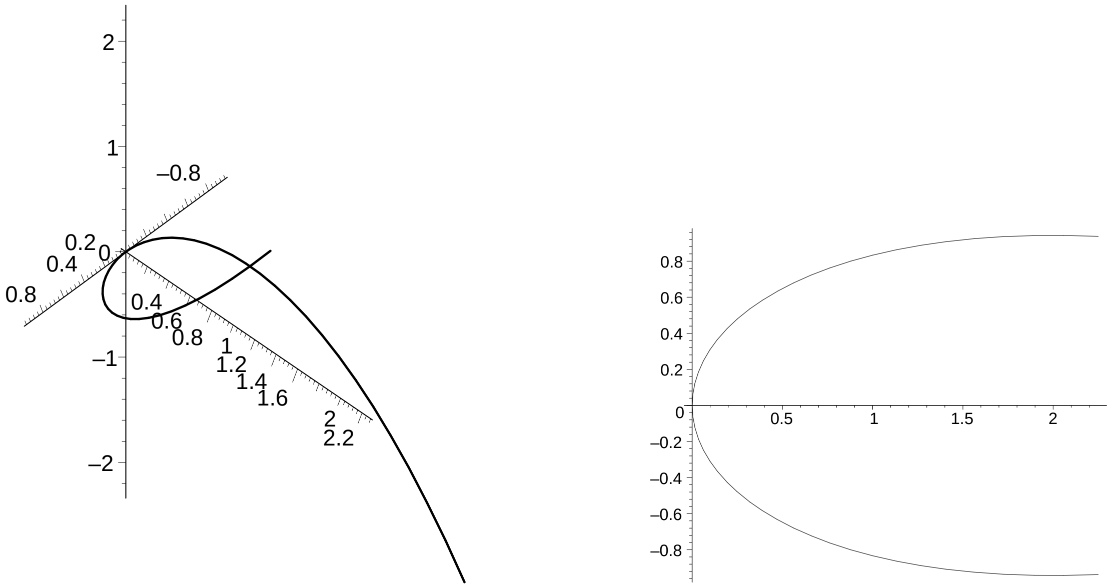


Figure 1.9: A space curve and its projection to the **nt**-plane

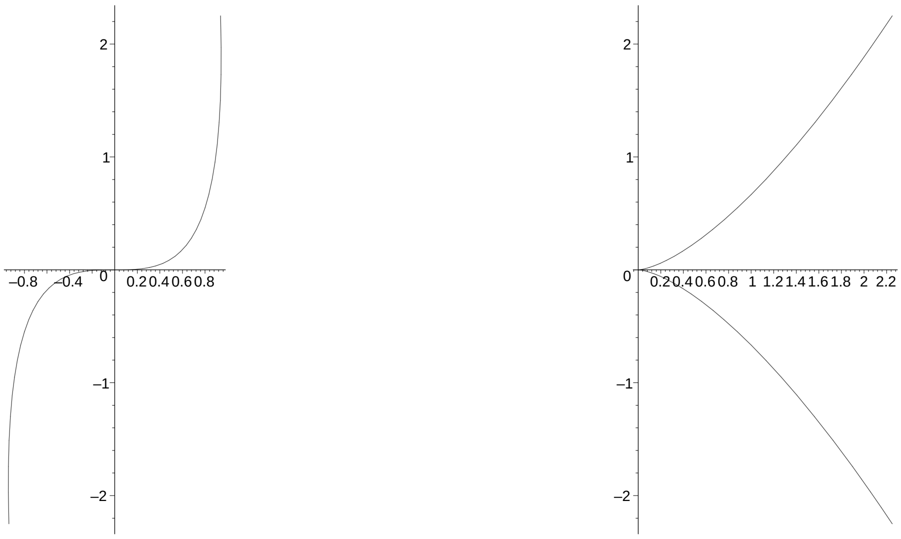


Figure 1.10: Projections to the **tb**-plane and to the **nb**-plane

3. What is the limit position of the circles through three points on a curve at an inflection point?

4. Show that a space curve has constant positive curvature *and* constant torsion if and only if it is a helix.
   (Use Thm. 1.74).

5. A space curve C is called a *generalised helix* if there is a unit vector $\mathbf{u} \in \mathbf{R}^3$ (its axis) such that the angle θ between tangent lines along the curve C and $\mathbf{u}$ is constant. Show:

   (a) A helix is a generalised helix.

   (b) A curve with parametrization $\mathbf{r}(t) = [x(t), y(t), z(t)]$ such that $y'(t)^2 = 2x'(t)z'(t)$ for all t is a generalised helix with axis $\mathbf{u} = [1, 0, 1]$.

   (c) A space curve without inflection points is a generalised helix if and only if the quotient $\frac{\tau}{\kappa}$ is constant.

   Hints: Let $\{\mathbf{t}, \mathbf{n}, \mathbf{b}\}$ denote the moving frame along C. Show for a generalised helix without inflection points:

   - $\mathbf{n} \cdot \mathbf{u} = \mathbf{0}$.
   - $\mathbf{u} = \cos\theta \mathbf{t} + \sin\theta \mathbf{b}$
   - $\cot\theta = \frac{\tau}{\kappa}$.

   Reverse this reasoning for a curve with constant $\frac{\tau}{\kappa}$ using appropriate *definitions* of θ and $\mathbf{u}$.

1.5 Global properties

Until now, we have focussed on *local* properties of curves, determined at a point by an arbitrarily small portion of the curve around it. The following section deals with global geometric properties of *closed* smooth curves having a periodic parametrization $\mathbf{r} : \mathbf{R} \to \mathbf{R}^3$ with period L, i.e., $\mathbf{r}(t + L) = \mathbf{r}(t)$. (Of course, one may restrict attention to a closed interval of size L as the domain. But we need also the (coinciding!) derivatives at the ends of such an interval.) Most results in this section are given without proofs.

1.5.1 The four-vertex theorem for plane curves

Definition 1.78 *Let C denote a plane smooth closed curve.*

1. *C is called an* oval *if its oriented curvature is either non-negative or non-positive everywhere, i.e., if it turns either clockwise or anti-clockwise.*
2. *A point P_0 on C is called a* vertex *if one has for some (and thus any!) smooth parametrization $\mathbf{r}: \mathbf{R} \to \mathbf{R}^2$ of C with $\mathbf{r}(t_0) = \overrightarrow{OP_0}$: $\kappa'(t_0) = 0$.*

It is clear, that the differentiable curvature function κ attains both a maximal and a minimal value within a period interval. Hence, an oval must have at least two vertices.

Example 1.79 *An ellipse is an oval. Remark that its curvature has two maxima and two minima (within a period), cf. Fig. 1.4.*

The following much more general result is true:

Theorem 1.80 ***(Four vertex theorem)*** *Let C denote an oval with smooth parametrization $\mathbf{r}: \mathbf{R} \to \mathbf{R}^2$ and period L. Then C has at least four* vertices, *i.e., for every $a \in \mathbf{R}$, there exist $a \leq t_1 < t_2 < t_3 < t_4 < a+L$ such that $\kappa'(s) = 0$.*

Illustration 1.81 *closed curves and their vertices, ovals and non-ovals.*

1.5.2 Total curvature and Fenchel's theorem

Definition 1.82 *Let C denote a closed smooth curve with unit speed parametrization $\mathbf{r}$ of period L. Its* total curvature *is defined as*

$$K(C) = \int_a^{a+L} \kappa \, ds.$$

For a *plane curve*– using oriented curvature as in Sect. 1.35 – there is a clean characterization of the total curvature: Let $\theta(a)$ denote the angle between $\mathbf{i} = [0,1] \in \mathbf{R}^2$ and $\mathbf{t}(a) = \mathbf{r}'(a)$. There is a *continuous* angle function $\theta : \mathbf{R} \to \mathbf{R}$ extending $\theta(a)$ such that $\theta(s)$ is the angle between $\mathbf{i}$ and $\mathbf{t}(s)$. (Remark that $\theta(s)$ is only well-determined up to integer multiples of 2π; the continuity of θ requires to do a particular choice once $\theta(a)$ has been selected.) The *integer* $\frac{\theta(a+L)-\theta(a)}{2\pi}$ is called the *winding number* $w(C)$ of the curve C (with period L).

Proposition 1.83

$$K(C) = 2\pi w(C).$$

In particular, $w(C)$ is well-determined, i.e., independent of the choice of a.

Proof: In Sect. 1.3.3, we got for a plane curve with unit speed parametrization, that oriented curvature satisfies: $\kappa(s) = \theta'(s)$. We can apply the fundamental theorem in Calculus to yield:

$$K(C) = \int_a^{a+L} \kappa ds = [\theta(s)]_a^{a+L} = \theta(a+L) - \theta(a).$$

□

Illustration 1.84 *closed curves, turning tangent, angle function, total curvature, winding number*

The total curvatures of space curves are usually less rigid. Let us nevertheless report on the following results:

Theorem 1.85 ***(Fenchel's theorem)*** *The total curvature of every closed smooth curve C satisfies:*

$$K(C) \geq 2\pi.$$

Equality holds if and only if C is an oval *plane curve (in some plane in $\mathbf{R}^3$).*

(This result does not contradict Prop. 1.83: for a plane non-oval curve, oriented curvature switches sign, while the curvature of the same curve regarded as a space curve has constant sign and thus greater total curvature!)

Total curvature for space curves carries even more information. A space curve is called *unknotted* if it can be continuously deformed into the plane unit circle and *knotted* else.

Theorem 1.86 ***(Fary-Milnor theorem)*** *For a knotted closed curve C, the total curvature satisfies:*

$$K(C) \geq 4\pi.$$

Illustration 1.87 *closed unknotted and notted curves and their total curvatures.*

1.5.3 Exercises

1. Determine the total curvature of the plane curve "figure 8" (with a double point).

2 Surfaces in space

2.1 Surfaces in space – definitions and examples

2.1.1 Parametrized and regular surfaces

Definitions and their interpreation

We saw in the last chapters how to make a precise formalism covering the idea of a curve, i.e., a line which one bends and twists in plane or space. Here we want to make precise what we mean by a surface in space. To start with, one could think of this as a plane which is then bent, twisted and stretched at will in 3-space, but we have to be more precise.

If we imitate the definition of a curve, a surface should be the image of a function $f(u,v) = (f_1(u,v), f_2(u,v), f_3(u,v))$ of two variables, where we would then have to impose suitable restrictions on the f_i.

The differentiability is easy to generalize, but what should we ask for to get the equivalent of $\alpha'(t) \neq 0$? Remember the derivative of a function of several variables:

Definition 2.1 *Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^k$ be a differentiable function of n variables. The* differential df_q *of f at $q \in U$ is the linear map from $\mathbf{R}^n$ to $\mathbf{R}^k$ given by the Jacobi matrix Df_q of f at q:*

$$df_q(v_1,\ldots,v_n) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(q) & \frac{\partial f_1}{\partial x_2}(q) & \ldots & \frac{\partial f_1}{\partial x_n}(q) \\ \frac{\partial f_2}{\partial x_1}(q) & \frac{\partial f_2}{\partial x_2}(q) & \ldots & \frac{\partial f_2}{\partial x_n}(q) \\ \vdots & \ddots & \vdots & \\ \frac{\partial f_k}{\partial x_1}(q) & \frac{\partial f_k}{\partial x_2}(q) & \ldots & \frac{\partial f_k}{\partial x_n}(q) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

It turns out that the proper equivalent to the requirement $\alpha'(t) \neq 0$ in the case of a function $f(u,v)$ of two variables into $\mathbf{R}^3$ is to require *injectivity* of the linear map $Df_q : \mathbf{R}^2 \to \mathbf{R}^3$ at every point q in the domain of f.

This approach is taken at several places in the history of differential geometry and it leads to

Definition 2.2 *A* parametrized *surface is a map $f : U \to \mathbf{R}^3$, where $U \subseteq \mathbf{R}^2$ is an open subset, f is a smooth map and the differential $df_{(u,v)} : \mathbf{R}^2 \to \mathbf{R}^3$ is injective for all $(u,v) \in U$.*

Example 2.3 *Examples of such surfaces are in Fig. 2.1. We give parametrizations of them here:*

- $U = \mathbf{R}^2$, $f(u,v) = (u, v, u^2 + v^2)$
- $U = \{(u,v) \in \mathbf{R}^2 | u \neq 0 \text{ and } u \neq 1\}$

$$f(u,v) = \begin{cases} (u+1, v, 0) & \text{for } u < 0 \\ (0, v, u) & \text{for } 0 < u < 1 \\ (0, v, 1-u) & \text{foru} > 1 \end{cases}$$

- $U = \mathbf{R}^2$, $f(u,v) = (\mu_1(u), \mu_2(u), v)$ *where* $\mu(u) = (1, u)$ *for* $u \leq 0$ *and* $\mu(u) = (\cos(2\pi e^{-\frac{1}{u^2}}), \sin(2\pi e^{-\frac{1}{u^2}}))$ *for* $u > 0$.

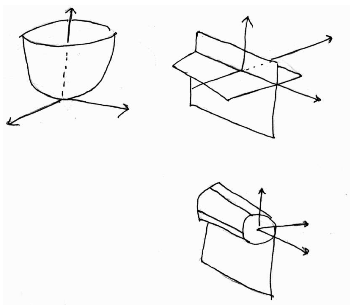


Figure 2.1: Parametrized surfaces. The last two are not regular surfaces

This is almost, but still not quite, the kind of surfaces we have in mind. We want a surface to be a subset of $\mathbf{R}^3$ which "looks like a plane" locally, i.e., for a small enough neighbourhood in $\mathbf{R}^3$ at each of its points. The last two examples above do not satisfy this requirement on the "intersection" line. No matter how small a neighbourhood we take, within $\mathbf{R}^3$, around $p = (0, v, 0)$ in Example 1.3.2 or around $p = (1, 0, v)$ in Example 1.3.3, the surface looks like the intersection of *two* planes in a line. To avoid such "singular" points, we will impose an additional requirement (3. below) that might look complicated to start with – until the reader gets used to it:

Definition 2.4 *A subset $S \subset \mathbf{R}^3$ is a regular surface if for any point $p \in S$ there is an open subset $V \subseteq \mathbf{R}^3$ containing p, an open subset $U \subseteq \mathbf{R}^2$ and a bijective map $\mathbf{x} : U \to S \cap V$ such that*

1. $\mathbf{x}$ *is* smooth*: If $\mathbf{x}(u, v) = (x_1(u, v), x_2(u, v), x_3(u, v))$, then each x_i has partial derivatives of all orders.*
2. $\mathbf{x}$ *is* regular*: For any point $(u, v) \in U$, the differential $d\mathbf{x}_{(u,v)} : \mathbf{R}^2 \to \mathbf{R}^3$ is injective.*
3. $\mathbf{x} : U \to S \cap V$ *is a* homeomorphism*: For any open subset $\tilde{U} \subset U$ there is an open subset $\tilde{V} \subseteq \mathbf{R}^3$ such that $\mathbf{x}(\tilde{U}) = \tilde{V} \cap S$.*

The map $\mathbf{x}$ *is called a* parametrization*, a* coordinate system *or a* coordinate chart *around p. The set $V \cap S$ is called a* coordinate neighborhood *of p.*

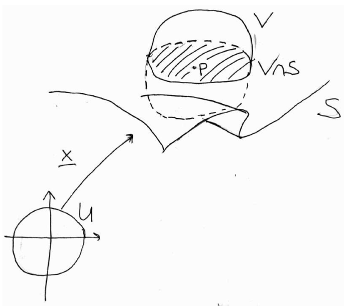


Figure 2.2: A regular surface and a coordinate system

Remember that a homeomorphism is a *bijective* continuous map with a *continuous inverse*. The open sets of S are, by definition, those of the form $\tilde{V} \cap S$ with $\tilde{V} \subset \mathbf{R}^3$ open. A map between $f : X \to Y$ between topological spaces (such as U and $V \cap S$) is continuous if $f^{-1}(Z) \subset X$ is continuous for every open subset $Z \subset Y$. The reader will check easily that the requirement in Definition 2.4.3. is equivalent to the continuity of the inverse $\mathbf{x}^{-1}$ of the parametrization $\mathbf{x}$. You should perhaps also investigate why this requirement is *not* satisfied in Examples 1.3.2 and 1.3.3 (or have a look at Example 2.8).

Condition 2 can be formulated in various ways, as the reader may recall from linear algebra. Here are a few options: The Jacobi matrix $D\mathbf{x}_q$ is defined as

$$D\mathbf{x}_q = \begin{pmatrix} \frac{\partial x_1}{\partial u}(q) & \frac{\partial x_1}{\partial v}(q) \\ \frac{\partial x_2}{\partial u}(q) & \frac{\partial x_2}{\partial v}(q) \\ \frac{\partial x_3}{\partial u}(q) & \frac{\partial x_3}{\partial v}(q) \end{pmatrix}.$$

Tence the following are equivalent:

- $D\mathbf{x}_q$ is injective.
- The two column vectors of $D\mathbf{x}_q$ are linearly independent.
- At least one pair of row vectors of $D\mathbf{x}_q$ is linearly independent.
- At least one of the 2-by-2 submatrices

$$\begin{pmatrix} \frac{\partial x_1}{\partial u}(q) & \frac{\partial x_1}{\partial v}(q) \\ \frac{\partial x_2}{\partial u}(q) & \frac{\partial x_2}{\partial v}(q) \end{pmatrix}, \begin{pmatrix} \frac{\partial x_1}{\partial u}(q) & \frac{\partial x_1}{\partial v}(q) \\ \frac{\partial x_3}{\partial u}(q) & \frac{\partial x_3}{\partial v}(q) \end{pmatrix}, \begin{pmatrix} \frac{\partial x_2}{\partial u}(q) & \frac{\partial x_2}{\partial v}(q) \\ \frac{\partial x_3}{\partial u}(q) & \frac{\partial x_3}{\partial v}(q) \end{pmatrix}$$

  has non-zero determinant.

Lemma 2.5 *With notation from above, let $\mathbf{x} : U \to V \cap S$ be a continuous bijection and suppose there is a continuous map $F : V \to \mathbf{R}^2$ such that $F_{|V\cap S} = \mathbf{x}^{-1}$, then $\mathbf{x}$ satisfies 2.4.3*

Proof: Let $\tilde{U} \subseteq U$ be an open subset. Since F is continuous, $\tilde{V} = F^{-1}(\tilde{U})$ is open in $\mathbf{R}^3$. We have to see that $\mathbf{x}(\tilde{U}) = \tilde{V} \cap S$

$\subseteq$ Let $q \in \tilde{U}$. Then $q = F(\mathbf{x}(q))$, so $\mathbf{x}(q) \in F^{-1}(\tilde{U}) = \tilde{V}$ and clearly $\mathbf{x}(q) \in S$.

$\supseteq$ Let $r \in \tilde{V} \cap S$. Then $\mathbf{x}^{-1}(r) = F(r) \in \tilde{U}$, and $r = \mathbf{x}(\mathbf{x}^{-1}(r))$.

□

To give the reader another picture of what 2.4.3 means, we state the following lemma:

Lemma 2.6 *Let $\mathbf{x} : U \to V \cap S$ be a coordinate system on a regular surface S. Suppose $(q_n)_{n\in\mathbb{N}}$ in $\mathbf{x}(U) \subseteq S \cap V$ converges (as a sequence in $\mathbf{R}^3$) to $q \in S \cap V$, then the sequence $(\mathbf{x}^{-1}(q_n))_{n\in\mathbb{N}}$ converges to $\mathbf{x}^{-1}(q)$ in U.*

Proof: Let $\tilde{U} \subset U$ be a neighborhood of $\mathbf{x}^{-1}(q)$ and $\mathbf{x}(\tilde{U}) = \tilde{V} \cap S$, where $\tilde{V} \subset \mathbf{R}^3$ is open. Then there is an N such that $n \geq N \Rightarrow q_n \in \tilde{V}$ and hence $\mathbf{x}^{-1}(q_n) \in \tilde{U}$.

□

Remark 2.7 *If you are used to defining convergence in terms of open balls, the above proof could be stated as follows: Let $B(\mathbf{x}^{-1}(q), \varepsilon) \subset U$, then there is an open subset $\tilde{V} \in \mathbf{R}^3$ s.t. $\mathbf{x}(B(\mathbf{x}^{-1}(q), \varepsilon)) = \tilde{V} \cap S$. Let $B(q, \delta) \subset \tilde{V}$. Then there is an N such that $n \geq N \Rightarrow q_n \in B(q, \delta)$ and hence $\mathbf{x}^{-1}(q_n) \in B(\mathbf{x}^{-1}(q), \varepsilon)$.*

Example 2.8 *The last two parametrized surfaces in Ex. 2.3 are* not *regular surfaces; they violate Lem. 2.6; the reader can check this by constructing suitable convergent sequences on the surfaces. In particular they violate Def. 2.4 condition 3.*

2.2 Particular classes of regular surfaces

2.2.1 Graph surfaces

Proposition 2.9 *Let $U \subset \mathbf{R}^2$ denote an open set and $f : U \to \mathbf{R}$ be a smooth function. The graph of f is defined as $G = \{[u, v, f(u, v)] \in \mathbf{R}^3 | \ (u, v) \in U\} \subset \mathbf{R}^3$. It is a regular surface covered by the single parametrization*

$$\mathbf{x} : U \to \mathbf{R}^3, \quad (u, v) \mapsto [u, v, f(u, v)].$$

Proof: We have to see that $\mathbf{x}$ is in fact a parametrization. The map $\mathbf{x}$ is clearly smooth, and it is easy to check, that it is a bijection from U onto G. The Jacobian matrix for the differential of $\mathbf{x}$ at $(u_0, v_0) \in U$ is

$$D\mathbf{x}_{(u_0,v_0)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial u}(u_0, v_0) & \frac{\partial f}{\partial v}(u_0, v_0) \end{pmatrix}$$

Condition 2 is satisfied as well since the two top rows are clearly linearly independent.

For the last condition, let $\pi : \mathbf{R}^3 \to \mathbf{R}^2$ be the map $\pi(x, y, z) = (x, y)$. This is a linear map, hence continuous; and $\pi_{|G} = \mathbf{x}^{-1}$.

□

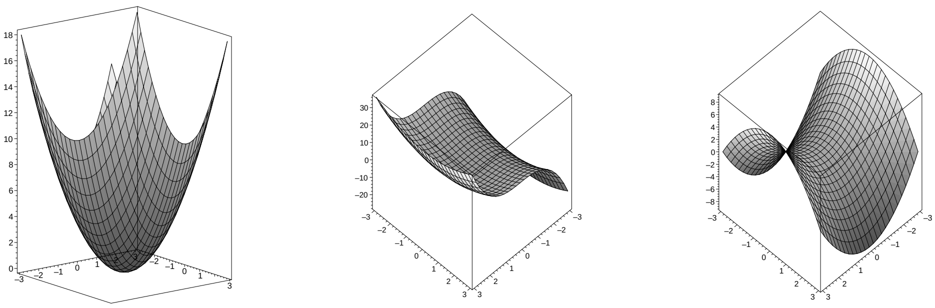


Figure 2.3: The graphs of $(u^2 + v^2)$, $(u^3 + v^2)$ and $(u^2 - v^2)$.

The sphere with the 6 graph coordinate systems.

Figure 2.4: The 2-sphere covered by six coordinate systems

Corollary 2.10 *Let U and f be as above, then the sets consisting of all points in $\mathbf{R}^3$ of the form $[u, v, f(u, v)]$, $[u, f(u, v), v]$, resp. $[f(u, v), u, v]$ are regular surfaces. They are all called graphs of f.*

Example 2.11 *This already provides us with lots of regular surfaces, some of which may be seen in Fig. 2.3.*

Definition 2.12 *When $U \subset \mathbf{R}^2$ is open and $f : U \to \mathbf{R}$ is smooth, a parametrization $\mathbf{x}$ of the kind*

$$(u, v) \mapsto [u, v, f(u, v)] \quad (u, v) \mapsto [u, f(u, v), v] \quad (u, v) \mapsto [f(u, v), u, v] \qquad (2.1)$$

will be called a graph coordinate system *for $\mathbf{x}(U)$.*

Example 2.13 *Let Π be a plane in $\mathbf{R}^3$. Then Π is a regular surface. The proof is left as an exercise: construct a graph coordinate system!*

Example 2.14 *The sphere $S^2 = \{(x, y, z) \in \mathbf{R}^3 | x^2 + y^2 + z^2 = 1\}$ in Fig. 2.4 is a regular surface: Let $U = \{(u, v) \in \mathbf{R}^2 | u^2 + v^2 < 1\}$. We define a set of 6 parametrizations $\mathbf{x}_i : U \to V_i \cap \mathbf{R}^3$ covering all of S^2 as follows:*

$$x_1(u, v) = (u, v, \sqrt{1 - (u^2 + v^2)}),\ V_1 = \{(x_1, x_2, x_3) \in \mathbf{R}^3 | x_3 > 0\}$$

$$x_2(u, v) = (u, v, -\sqrt{1 - (u^2 + v^2)}),\ V_2 = \{(x_1, x_2, x_3) \in \mathbf{R}^3 | x_3 < 0\}$$

$$x_3(u, v) = (u, \sqrt{1 - (u^2 + v^2)}, v),\ V_3 = \{(x_1, x_2, x_3) \in \mathbf{R}^3 | x_2 > 0\}$$

$$x_4(u,v) = (u, -\sqrt{1-(u^2+v^2)}, v),\ V_4 = \{(x_1,x_2,x_3) \in \mathbf{R}^3 | x_2 < 0\}$$
$$x_5(u,v) = (\sqrt{1-(u^2+v^2)}, u, v),\ V_5 = \{(x_1,x_2,x_3) \in \mathbf{R}^3 | x_1 > 0\}$$
$$x_6(u,v) = (-\sqrt{1-(u^2+v^2)}, u, v),\ V_6 = \{(x_1,x_2,x_3) \in \mathbf{R}^3 | x_1 < 0\}$$

These are all graph coordinate systems, and hence they satisfy the conditions in Def. 2.4. Please check that every point $p \in S^2$ is covered by at least one of the six coordinate systems.

2.2.2 Surfaces of rotation

Let $f :]a,b[\to \mathbf{R}$ be a smooth map with $f(t) > 0$ for all $t \in]a,b[$. Then the subset of $\mathbf{R}^3$ given by

$$S = \{(f(u)\cos(v), f(u)sin(v), u) |\ a < u < b, 0 < v \leq 2\pi\}$$

is a regular surface.

Interpretation: The intersection of S with the plane $Z = u, a < u < b$, is a circle with radius $f(u)$ with center at $[0,0,u]$ on the Z-axis. Hence the surface arises by rotating the regular(!) curve consisting of points $[f(u), 0, u]$ around the Z-axis.

To prove that S is a regular surface, we have to provide local coordinates around all points. Let $U_1 =]a,b[\times]0,2\pi[$ and $U_2 =]a,b[\times]-\pi,\pi[$ and let $\mathbf{x}_i(u,v) = (f(u)\cos(v), f(u)sin(v), u)$. We claim that this defines two coordinate charts covering S. Differentiability is clear. For $\mathbf{x}_1$, let

$$V_1 = \mathbf{R}^3 \setminus \{(x,y,z) \in \mathbf{R}^3 | y = 0 \text{ and } x \geq 0\}.$$

We claim that $\mathbf{x}_1$ is a bijection to $V_1 \cap S$. For a proof, we describe its inverse: $F_1(x,y,z) = (z, \arccos(\frac{x}{\sqrt{x^2+y^2}}))$ for $y \geq 0$ and $F_1(x,y,z) = 2\pi - \arccos(\frac{x}{\sqrt{x^2+y^2}}))$ for $y < 0$. To see that $F_1 : V_1 \to \mathbf{R}^2$ is continuous, there might be a problem at $y = 0$. From the definition of V_1, we know that $x < 0$, and hence $\arccos(\frac{x}{\sqrt{x^2}})) = \arccos(-1) = \pi = 2\pi - \arccos(\frac{x}{\sqrt{x^2}}))$.

We leave it to the reader to see that the restriction of F to $S \cap V_1$ is an inverse to $\mathbf{x}_1$, which is then a bijection $\mathbf{x}_1 : U_1 \to S \cap V_1$.

Regularity: We calculate the differential

$$(D\mathbf{x}_1)_{[u,v]} = \begin{pmatrix} f'(u)\cos(v) & -f(u)\sin(v) \\ f'(u)sin(v) & f(u)\cos(v) \\ 1 & 0 \end{pmatrix}$$

torus

Figure 2.5: A torus

The 2-by-2 subdeterminants are $f'(u)f(u)$, $f(u)\sin(v)$ and $-f(u)\cos(v)$. Since $f(u) > 0$, at least one of the latter two is non-zero.

To see that $\mathbf{x}_2$ is a coordinate chart, imitate the above arguments with $V_2 = \mathbf{R}^3 \setminus \{(x,y,z)|y = 0 \text{ and } x \leq 0\}$ and a proper choice of $F_2 : V_2 \to \mathbf{R}^2$.

Remark 2.15 *Surfaces of rotation may be defined for more general plane curves. Let a curve be given with a regular parametrization $I\mathbf{R}^3, u \mapsto [f(u), 0, g(u)]$ with $g(u) > 0$ (right half-plane!) and $f'(u)^2 + g'(u)^2 < 0$ for all u in a closed interval I. In order to avoid self-intersections of the surface, we have to avoid self-interesections of the original curve, i.e., the parametrization needs to be injective, apart from perhaps at the end points of the interval. If this is the case, then the following formula can be used for parametrizations of the surface of rotation (you might have to restrict to the interior of the interval I or otherwise take particular care at the end points of the interval):*

$$(u,v) \mapsto [f(u)\cos v, f(u)\sin v, g(u),\ u \in I, v \in [0, 2\pi].$$

We leave it as an exercise to make these requirements precise.

Example 2.16 *A torus $T \subset \mathbf{R}^3$ arises as surface of rotation from a circle of radius $a > 0$ in the XZ-plane with center at $[b, 0, 0]$ with $a < b$, cf. Figure 2.5. Hence such a torus is a regular surface. Parametrizations $\mathbf{x}_i$, $i = 1, 2$ are of the form $\mathbf{x}_i(u,v) = [(b + a\cos u)\cos v, (b + a\cos u)\sin v, a\sin u]$.*

2.2.3 Ruled surfaces

2.3 Implicitly given surfaces

Like planes can be defined as the set of solutions of linear equations in three variables, many surfaces have a description as the set of solutions of (in general non-linear) equations – of a slightly restricted kind.

Example 2.17 *1. The solutions of the equation $X^2 + Y^2 + Z^2 = r^2$ describe, for every $r \neq 0$, a sphere around the origin with radius r.*

2. *Consider the equation $X^2 + Y^2 - Z^2 = a$. The set of solutions consists of infinitely many circles in planes $Z = c$, since then $r^2 = X^2 + Y^2 = c^2 + a$. This equation has a solution for $c^2 + a \geq 0$; always for $a \geq 0$, but only for $|c| \geq \sqrt{-a}$ for negative a. For $a = 0$, the set of solutions is a* cone *with a singular point at the origin (a circle of "radius 0"); this set is* not *a regular surface. For all values $a \neq 0$, the set of solutions of the equation above is in fact a regular surface; for an illustrations, check Figure 2.7. The surface in the middle represents a conic that solves the equation above for $a > 0$; the surface on the right side arises as the set of solutions for $a < 0$.*

What makes the distinction between equations with regular surfaces as their set of solutions and those that have not? This is what we will investigate now:

Definition 2.18 *Let $X \subset \mathbf{R}^3$ be an open subset and let $f : X \to \mathbf{R}$ be a* smooth *function.*

1. *A point $\mathbf{p} \in X$ is a* regular point *for f if the gradient vector $\nabla f(p) = [\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial z}(p)] \neq \mathbf{0}$.*
2. *A number $a \in \mathbf{R}$ is a* regular value *for f if* every *solution $\mathbf{p}$ of the equation $f(\mathbf{p}) = a$ is regular.*

Example 2.19 *Let $f : \mathbf{R}^3 \to \mathbf{R}$ be given by $f(x, y, z) = x^2 + y^2 + z^2$.*
Claims:

1. *Every point $\mathbf{p} \in \mathbf{R}^3, \mathbf{p} \neq \mathbf{0}$, is a regular point.*
2. *Every real number $a \neq 0$ is a regular value.*

To see this, we calculate the gradient $\nabla f(x_0, y_0, z_0)) = (2x_0, 2y_0, 2z_0)$ which is only equal to $\mathbf{0}$ for $x_0 = y_0 = z_0 = 0$. Hence $a \in \mathbf{R}$ is a regular value if and only if $f(0, 0, 0) \neq a$, i.e., if and only if $a \neq 0$.

The importance of the terms regular point and regular value stem from a particular case of the implicit function theorem, to wit:

Theorem 2.20 *If $X \subseteq \mathbf{R}^3$ is open, $f : X \to \mathbf{R}$ is a smooth function and $a \in \mathbf{R}$ is a regular value for f, then $S := f^{-1}(a) = \{(x, y, z) \in X | f(x, y, z) = a\}$ is a regular surface.*

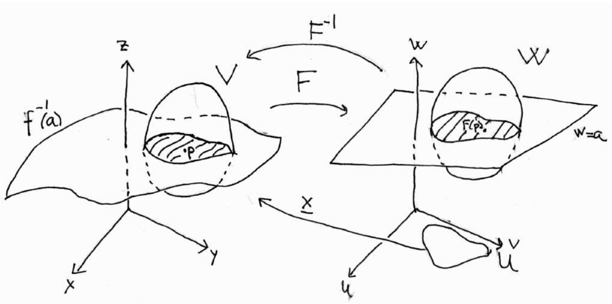


Figure 2.6: Illustration for thm. 2.20

The requirement in Theorem 2.20 can also be stated as: All elements of the set of solutions $S := f^{-1}(a)$ are regular points, i.e., for every $\mathbf{p} = (x, y, z) \in X$ with $f(x, y, z) = a$, at least one of the partial derivatives $\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial z}(p)$ does not vanish.

Here is the version of the implicit function theorem that you find most often in texbooks in mathematical analysis:

Theorem 2.21 *Assume $X \subseteq \mathbf{R}^3$ to be open, $f : X \to \mathbf{R}$ to be a smooth function, $a \in \mathbf{R}$, and $\mathbf{p}_0 = (x_0, y_0, z_0) \in X$ such that $f(\mathbf{p}_0) = a$ and $\frac{\partial f}{\partial z}(\mathbf{p}_0) \neq 0$. Then there exists an open neighbourhood U of (x_0, y_0) in $\mathbf{R}^2$, an open interval I containing z_0 and a smooth function $g : U \to I$ such that $g(x_0, y_0) = z_0$ and $f(x, y, g(x, y)) = a$. Moreover, $\{(x, y, z) \in U \times I | \, f(x, y, z) = a\} = \{(x, y, g(x, y)) \, (x, y) \in U\}$.*

That theorem tells us thus that the graph of the function g consists of solutions of the equation $f(\mathbf{p}) = a$ and that it contains the point $\mathbf{p}_0$. Note that U can only contain points (x, y) such that there exists z with $(x, y, z) \in U$ ("over U"). Remark also that the theorem only states the *existence* of such a function g. In general, there is no way to give a description of g by way of a closed formula of simple terms.

To prove Theorem 2.20 using the implicit function theorem (as stated in Theorem 2.21), note first that we can suppose, without loss of generality, that $\frac{\partial f}{\partial z}(\mathbf{p}_0) \neq 0$. (If, instead, $\frac{\partial f}{\partial x}(\mathbf{p}_0) \neq 0$, we conclude the existence of a smooth function $h : U \to \mathbf{R}$ such that $h(y_0, z_0) = x_0$ and $f(h(y, z), y, z) = a$.)

In any case, we may conclude that there is a neighbourhood W of $\mathbf{p}_0 \in \mathbf{R}^3$ such that the set $S \cap W$ consisting of the solutions of the equation $f(\mathbf{p}) = a$ can be parametrized as *graph* of a smooth function. Apply Corollary 2.10.

Remark 2.22 *The reader may have seen a statement of the implicit function theorem that is more general that the one described in Theorem 2.21 for a real* *function of three variables. The generalized version deals with smooth (vector) function $f : X \to \mathbf{R}^m$ for an open domain $X \subset \mathbf{R}^{n+m}$. If $f(\mathbf{p}) = \mathbf{a}$ for a point $\mathbf{p} = (\mathbf{p}_n, \mathbf{p}_m)) \in X$ and if a certain $m \times m$ minor of the Jacobian $Df_{\mathbf{p}}$ does not vanish, then there exist neighbourhoods $U \subset \mathbf{R}^n$ of $\mathbf{p}_n$ and $V \subset \mathbf{R}^m$ of $\mathbf{p}_m$ and a smooth vector function $g : U \to V$ such that $g(\mathbf{p}_n) = \mathbf{p}_m$, $f(\mathbf{x}, g(\mathbf{x})) = \mathbf{a}$ and such that $\{(\mathbf{x}, \mathbf{y}) \in U \times V \; f(\mathbf{x}, \mathbf{y}) = \mathbf{a}\} = \{\mathbf{x}, g(\mathbf{x}) | \; \mathbf{x} \in U\}$.*

For the sake of completeness, we include the proof of Theorem 2.20 taking departure in the *inverse function theorem* that is usually proved in a textbook in mathematical analysis. The implicit function theorem is then usually deduced as a consequence; it is in fact equivalent to the inverse function theorem. We state the inverse function theorem without proof as

Theorem 2.23 *Let $X \subseteq \mathbf{R}^n$ be an open set and let $F : X \to \mathbf{R}^n$ be a smooth map. Suppose $\mathbf{p} \in X$ and $dF_{\mathbf{p}} : \mathbf{R}^n \to \mathbf{R}^n$ is a bijection. Then F has a local smooth inverse: There is an open set $V \subseteq X$ containing $\mathbf{p}$, and an open set $W \subseteq \mathbf{R}^n$ containing $F(\mathbf{p})$ such that $F : V \to W$ is a bijection and the inverse $F^{-1} : W \to V$ is smooth.*

Proof:[of Theorem 2.20] Let $\mathbf{p} = (x_0, y_0, z_0) \in f^{-1}(a)$. *We want to* find a parametrization of a neighbourhood $V \cap f^{-1}(a)$ of $\mathbf{p}$. The strategy is to use the inverse function theorem 2.23 on a suitably defined function $F : V \to \mathbf{R}^3$.

Suppose, perhaps after renaming the axis, that $\frac{\partial f}{\partial z}(\mathbf{p}) \neq 0$. Let $F : X \to \mathbf{R}^3$ be defined by $F(x, y, z) = (x, y, f(x, y, z))$, then

$$DF_{\mathbf{p}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{\partial f}{\partial x}(\mathbf{p}) & \frac{\partial f}{\partial y}(\mathbf{p}) & \frac{\partial f}{\partial z}(\mathbf{p})) \end{pmatrix}.$$

The determinant of $dF_{\mathbf{p}}$ is non-zero, so $dF_{\mathbf{p}}$ is a bijection. By the inverse function theorem, there is an open set $V \subseteq X$ containing $\mathbf{p}$, an open set $W \subseteq \mathbf{R}^3$ containing $F(\mathbf{p}) = (x_0, y_0, a)$ and a smooth function $F^{-1} : W \to V$ which is an inverse for the restriction of F to V, i.e., such that

1. For $(x, y, z) \in V$, $F^{-1}(F(x, y, z)) = (x, y, z)$
2. For $(u, v, w) \in W$, $F(F^{-1}(u, v, w)) = (u, v, w)$.

We claim that $U = \{(u, v) \in \mathbf{R}^2 | (u, v, a) \in W\}$ and $\mathbf{x} : U \to \mathbf{R}^3$ given by $\mathbf{x}(u, v) = F^{-1}(u, v, a)$ is a parametrization of $V \cap f^{-1}(a)$.

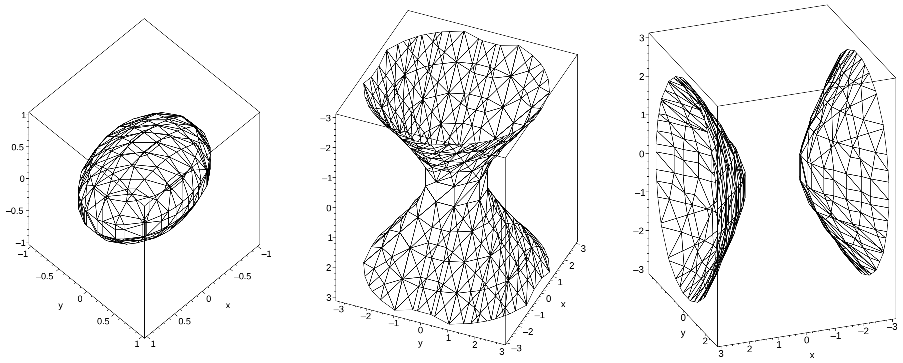


Figure 2.7: The quadratic surfaces from Ex. 2.24.

From 1):

$$(x, y, z) = F^{-1}(F(x, y, z)) = F^{-1}(x, y, f(x, y, z))$$

$$= (F_1^{-1}(x, y, f(x, y, z)), F_2^{-1}(x, y, f(x, y, z)), F_3^{-1}(x, y, f(x, y, z))),$$

we conclude that $F_1^{-1}(u, v, w) = u$ and $F_2^{-1}(u, v, w) = v$.

Hence $\mathbf{x}(u, v) = F^{-1}(u, v, a) = (u, v, F_3^{-1}(u, v, a))$ is a graph. We have to see that the image $\mathbf{x}(U) = V \cap f^{-1}(a)$.

From 2):

$$(u, v, w) = F(F^{-1}(u, v, w)) = F(u, v, F_3^{-1}(u, v, w)) = (u, v, f(F^{-1}(u, v, w))$$

Hence $\mathbf{x}(U) = F^{-1}(W \cap \{(u, v, a)\}) \subseteq f^{-1}(a) \cap V$: For $w = a$, $f(F^{-1}(u, v, a)) = F_3(F^{-1}(u, v, a)) = a$ and hence $F^{-1}(u, v, a) \in f^{-1}(a)$.

Moreover, $F^{-1}(W \cap \{(u, v, a)\}) \supseteq f^{-1}(a) \cap V$: If $q = (x, y, z) \in f^{-1}(a) \cap V$ then $F(q) = (x, y, a)$ and $F(q) \in W$, since $q \in V$.

Hence: F^{-1} is a bijection from $W \cap \{(u, v, a)\}$ to $f^{-1}(a) \cap V$.

$\square$

Example 2.24 *As a slight generalization of Example 2.17, we consider the quadratic surfaces in space that arise as the set S of solutions of the equation $f(x, y, z) = \frac{x^2}{a^2} \pm \frac{y^2}{b^2} \pm \frac{z^2}{c^2} = 1$ with a, b and c non-zero. Every such space S is a regular surface, since $\nabla f_{(x_0, y_0, z_0)} = (\frac{2x_0}{a^2}, \pm\frac{2y_0}{b^2}, \pm\frac{2z_0}{c^2}) \neq \mathbf{0}$ except for $x_0 = y_0 = z_0 = 0$. But $f(0, 0, 0) = 0$, and hence $(0, 0, 0) \notin S = f^{-1}(1)$ is not a solution.*
The three different types of regular quadratic surfaces are depicted in Fig.2.7

2.4 Coordinate systems

Coordinate systems for a regular surface are not unique, there are infinitely many of them. Using Thm. 2.20 we may have established that a subset $S \subset \mathbf{R}^3$ is a regular surface without knowing any coordinate patch explicitely.

The simplest situation arises for a graph S of a smooth function f. This subset $S \subset \mathbf{R}^3$ has a simple atlas consisting of only one coordinate patch. It is certainly not true, that all surfaces are graphs, but locally this does in fact hold! This will be proven in Prop. 2.27.

To begin with, we note another formulation of condition 3 from Definition 2.4: If we already know that S is a regular surface, we only have to check that a local parametrization satisfies conditions 1 and 2; condition 3 is for free.

Proposition 2.25 *Let $\mathbf{x} : U \to S$ be a coordinate system on a regular surface S and let $U_0 \subseteq U$ be any open subset. Then the restriction $\mathbf{x} : U_0 \to S$ is again a coordinate system on S.*

Proof: Left to the reader.

□

Actually, the result of Proposition 2.25 is equivalent to condition 3, which the reader may try to prove.

2.4.1 "Inverse" to a parametrization

What about the inverse of a parametrization $\mathbf{x} : U \to W \cap S$? The inverse $\mathbf{x}^{-1} : W \cap S \subset S$ into U has to be continuous. It makes no sense to ivestigate differentiability, since its domain $W \cap S$ is not an open subset of Euclidean space. Here is as close as we can come to a smooth "inverse":

Proposition 2.26 *Let $\mathbf{x} : U \to \mathbf{R}^3$ denote a parameterized surface. For every $p \in U \subset \mathbf{R}^2$ there exists an open subset $U_0 \subset U$ containing p such that the restriction $\mathbf{x}_0 = \mathbf{x}|_{U_o} : U_0 \to \mathbf{R}^3$ has a smooth "inverse" in the following sense: There exists an open set $V_0 \subset \mathbf{R}^3$ and a smooth map $F_0 : V_0 \to U_0$ such that $(F_0 \circ \mathbf{x}_0)(u, v) = (u, v)$ for every $(u, v) \in U_0$.*

Proof: The proof contains two steps:

Step 1 Consider the projections $\pi_i : \mathbf{R}^3 \to \mathbf{R}^2,\ i = 1, 2, 3$, leaving out the i-th coordinate and the smooth(!) compositions $\pi_i \circ \mathbf{x} : U \to \mathbf{R}^2$. Their Jacobi matrices $D(\pi_i \circ \mathbf{x})_p$ at p are 2×2-matrices obtained by deleting one of the rows of $D\mathbf{x}_\mathbf{p}$. Without loss of generality, we may assume that $D(\pi_3 \circ \mathbf{x})_\mathbf{p}$ has rank 2 and hence is invertible.

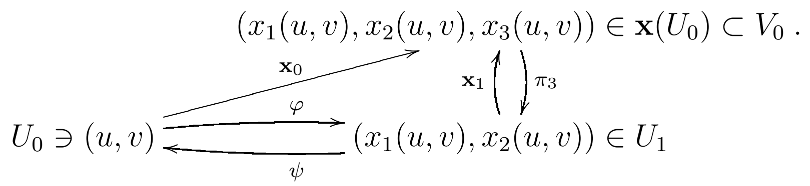
Applying the inverse function theorem (Thm. 2.23) to the map $\pi_3 \circ \mathbf{x}$, we conclude that there is an open set $U_0 \subset U$ containing p such that $U_1 := (\pi_3 \circ \mathbf{x})(U_0) \subset \mathbf{R}^2$ is open and such that the restriction of the map $\varphi := \pi_3 \circ \mathbf{x} : U_0 \to U_1$ is a *diffeomorphism* between open planar sets.

Step 2 This means that the map $\varphi : U_0 \to U_1$ has a *smooth inverse* $\psi : U_1 \to U_0$. Let $V_0 := \pi_3^{-1}(U_1)$ – an open subset of $\mathbf{R}^3$ – and let $\mathbf{x_0} : U_0 \to \mathbf{R}^3$ denote the restriction of $\mathbf{x}$ to U_0. Then $\mathbf{x}_0(U_0) \subset V_0$ since $(\pi_3 \circ \mathbf{x}_0)(U_0) = U_1$. The map $F_0 := \psi \circ \pi_3|_{V_o} : V_0 \to U_0$ is smooth as composition of two smooth maps. Moreover:

$$F_0 \circ \mathbf{x}_0 = (\psi \circ \pi_3) \circ \mathbf{x}_0 = \psi \circ (\pi_3 \circ \mathbf{x}_0) = \psi \circ \varphi = id|_{U_0}.$$

□

A synopsis of the proof above and also that of Proposition 2.27– explaining almost everything very briefly – is given in the *commutative diagram*



2.4.2 Existence of simple coordinate charts

The construction above has a simple consequence showing that every regular surface can be given a graph coordinate system – locally:

Proposition 2.27 *Let S denote a regular surface and $q \in S$ an arbitrary point. Then there exists an open neighbourhood $V_0 \subset \mathbf{R}^3$ with $q \in V_0$ and such that $V_0 \cap S$ is the image of a graph coordinate system.*

Proof: The proof may be considered as a **Step 3** added to the proof of Proposition 2.26 above. We "reparametrize" the map $\mathbf{x}_0 : U_0 \to \mathbf{R}^3$ using the diffeomorphism ψ (inverse to φ) to end up with $\mathbf{x}_1 := \mathbf{x}_0 \circ \psi : U_1 \to \mathbf{R}^3$. This map is smooth, its differential has rank 2 at every point in U_1 (chain rule!) and $\mathbf{x}_1$ is a homeomorphism onto its image $\mathbf{x}_1(U_1) = \mathbf{x}_0(U_0)$ as the composition of two homeomorphisms. Finally

$$\pi_3 \circ \mathbf{x}_1 = \pi_3 \circ (\mathbf{x}_0 \circ \psi) = (\pi_3 \circ \mathbf{x}_0) \circ \psi = \varphi \circ \psi = id|U_1;$$

hence $\mathbf{x}_1(u_1, v_1) = (u_1, v_1, F(u_1, v_1))$ with $F = \pi^3 \circ \mathbf{x}_1$ smooth. (Here, $\pi^3 : \mathbf{R}^3 \to \mathbf{R}$ denotes the map $(x, y, z) \mapsto z$).

□

We will use this result to obtain simple *local* arguments later: We may always assume that a sufficiently small neighborhood of a point $p \in S$ is the image of a graph coordinate system.

The following proposition tells us that if for some reason we already know that a subset $S \subset \mathbf{R}^3$ is a regular surface, then we can check whether a map $\mathbf{x} : U \to S$ is a coordinate system without going through the third condition in Def. 2.4.

Proposition 2.28 *Let $S \subset \mathbf{R}^3$ be a regular surface. Suppose $U \subseteq \mathbf{R}^2$ is open and $\mathbf{x} : U \to S$ is injective and satisfies conditions 1 and 2 in Def. 2.4, then condition 3 is also satisfied.*

Proof: Let $q \in \tilde{U} \subseteq U$ where $\tilde{U}$ is open. We have to find an open set $\tilde{V} \subseteq \mathbf{R}^3$ such that $\mathbf{x}(\tilde{U}) = \tilde{V} \cap S$. First we note that it is enough to verify this property locally: Suppose that we can find, for every $q \in U$, an open subset $q \in U_q \subset U$ and an open subset $V_q \subset \mathbf{R}^3$ such that $x(U_q) = V_q \cap S$. Let $\tilde{V} = \cup_{q \in \tilde{U}} V_q$. Then

$$\mathbf{x}(\tilde{U}) = \mathbf{x}(\bigcup_{q \in \tilde{U}} U_q) = \bigcup_{q \in \tilde{U}} \mathbf{x}(U_q)) = \bigcup_{q \in \tilde{U}} (V_q \cap S) = (\bigcup_{q \in \tilde{U}} V_q) \cap S = \tilde{V} \cap S.$$

Now we can refer to the proof of Proposition 2.27: Given $q \in U$, we found $U_q := U_0 \subset U$ and $V_q := V_0 \subset \mathbf{R}^3$ such that $x(U_q) = V_q \cap S$.

□

2.4.3 Change of coordinates

Proposition 2.29 *Let $\mathbf{x} : U \to S$ and $\mathbf{y} : \tilde{U} \to S$ be coordinate systems around $p \in S$. Then the change of coordinates*

$$\mathbf{x}^{-1} \circ \mathbf{y} : \mathbf{y}^{-1}(M) \to \mathbf{x}^{-1}(M)$$

with $M = \mathbf{y}(\tilde{U}) \cap \mathbf{x}(U)$, $p \in M$, is a diffeomorphism.

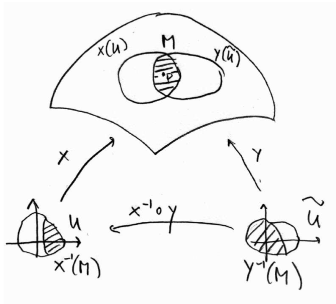


Figure 2.8: Change of coordinates

Proof: It is clear that $\mathbf{x}^{-1} \circ \mathbf{y} : \mathbf{y}^{-1}(M) \to \mathbf{x}^{-1}(M)$ is a bijection. Let $r \in \mathbf{y}^{-1}(M), q \in \mathbf{x}^{-1}(M)$ such that $p = \mathbf{y}(r) = \mathbf{x}(q)$; we will prove that $\mathbf{x}^{-1} \circ \mathbf{y}$ is differentiable at r. By Proposition 2.27, there exists an open subset $U_0 \subset \mathbf{x}^{-1}(M)$ with $q \in U_0$, an open subset $V_0 \subset \mathbf{R}^3$ and a smooth "inverse" map $F_0 : V_0 \to U_0$ such that $(F_0 \circ \mathbf{x})(u,v) = (u,v)$ for every $(u,v) \in U_0$. In particular, $\mathbf{x}^{-1}(x,y,z) = F_0(x,y,z)$ for every $(x,y,z) \in S \cap V_0$. Hence, $(\mathbf{x}^{-1} \circ \mathbf{y})(u,v) = (F_0 \circ \mathbf{y})(u,v)$ for all (u,v) in the open(!) set $\mathbf{y}^{-1}(\mathbf{x}(U_0))$ containing r. Withinn this set, the map $(\mathbf{x}^{-1} \circ \mathbf{y})$ coincides therefore with the smooth(!) map $F_0 \circ \mathbf{y}$.

To see that the inverse, $\mathbf{y}^{-1} \circ \mathbf{x}$ is differentiable at any point in $\mathbf{x}^{-1}(M)$ just interchange $\mathbf{y}$ and $\mathbf{x}$ in the above arguments.

□

2.5 Differentiable functions on surfaces.

Now it is time for calculus on surfaces. We want to study functions which are defined on a surface, but not necessarily only functions that arise as restriction of a function on $\mathbf{R}^3$. We will define when such functions are differentiable. In Chapter 3 we will define and investigate their differentials.

The following definition is basic:

Definition 2.30 *Let S be a regular surface and let $p \in S$.*

1. *A function $f : S \to \mathbf{R}$ is called* differentiable at p, *if there is a coordinate system $\mathbf{x} : U \to S$ around p such that $f \circ \mathbf{x} : \mathbf{x}^{-1}(\mathbf{x}(U) \cap V) \to \mathbf{R}$ is differentiable at $q = \mathbf{x}^{-1}(p)$. Moreover, f is* differentiable on S *– or just differentiable – if it is differentiable at every point in S.*

2. *Let S' denote a second regular surfaces. A function $\phi : S \to S'$ is called* differentiable at *p if there are coordinate systems $\mathbf{x} : U \to S$ around p and $\mathbf{y} : U' \to S'$ around $\phi(p)$ such that $\phi(\mathbf{x}(U)) \subset \mathbf{y}(U')$ and $\mathbf{y}^{-1} \circ \phi \circ \mathbf{x} : U \to U'$ is differentiable at $\mathbf{x}^{-1}(p)$. Moreover, ϕ is differentiable on S if it is differentiable at every point in S.*

3. *If ϕ is a smooth bijection with smooth inverse, then ϕ is called a* diffeomorphism*.*

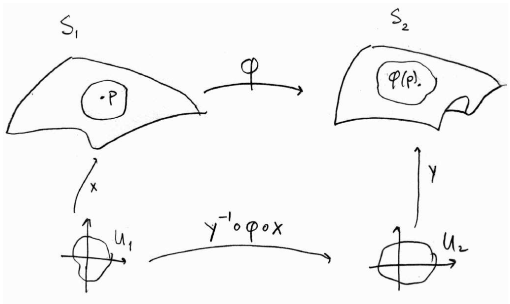


Figure 2.9: Differentiability

In Definition 2.30, we make use of (a) particular coordinate system(s) on the surface(s). Hence we need to check that the definition does not depend on this particular choice of coordinates. Our tool is Proposition 2.29 on change of coordinates. Below you find the proof for (2); the one for (1) is similar (and simpler).

Proposition 2.31 *Let $\mathbf{x}_1 : U_1 \to S$ and $\mathbf{x}_2 : U_2 \to S$ denote coordinate systems with $\mathbf{x}_1(q_1) = \mathbf{x}_2(q_2) = p$; likewise $\mathbf{y}_1 : U_1' \to S'$ and $\mathbf{y}_2 : U_2' \to S'$ denote coordinate systems covering $\phi(p)$. Then $\mathbf{y}_1^{-1} \circ \phi \circ \mathbf{x}_1$ is differentiable at q_1 if and only if $\mathbf{y}_2^{-1} \circ \phi \circ \mathbf{x}_2$ is differentiable at q_2.*

Proof: Assume that $\mathbf{y}_1^{-1} \circ \phi \circ \mathbf{x}_1$ is differentiable at q_1. Then

$$\mathbf{y}_2^{-1} \circ \phi \circ \mathbf{x}_2 = (\mathbf{y}_2^{-1}\mathbf{y}_1) \circ (\mathbf{y}_1^{-1} \circ \phi \circ \mathbf{x}_1) \circ (\mathbf{x}_1)^{-1} \circ \mathbf{x}_2, \tag{2.2}$$

and hence, by Proposition 2.29, $\mathbf{y}_2^{-1} \circ \phi \circ \mathbf{x}_2$ is a composition of differentiable maps ($q_2 \mapsto q_1 \mapsto \mathbf{y}_1^{-1}(\phi(p)) \mapsto \mathbf{y}_2^{-1}(\phi(p))$), and hence it is differentiable at q_2. Interchanging subscript 1 with subscript 2 yields a proof in the other direction.

The reader is invited to extend Figure 2.9 with additional arrows $\mathbf{x}_2$ and $\mathbf{y}_2$ and to follow (2.2) in the resulting diagram ("back and forth").

□

Corollary 2.32 *Let* $\mathbf{x} : U \to S$ *be a coordinate system on* S*, then* $\mathbf{x} : U \to \mathbf{x}(U)$ *is a diffeomorphism, when* U *is considered as the regular surface* $\{(u, v, 0) \in \mathbf{R}^3 | (u, v) \in U\}$.

Proof: An exercise.

□

2.6 Curves on a surface

The main tool used to investigate the local properties of a regular surface S are the curves that are contained in S (short: *curves on* S). More precisely, this is a curve with a (smooth) parametrization $\alpha : I \to \mathbf{R}^3$ such that $\alpha(t) \in S$ for all $t \in I$. How can one construct/describe such a curve?

This is quite easy locally, i.e., in presence of a parametrization $\mathbf{x} : U \to V \cap S$, $U \subset \mathbf{R}^2, V \subset \mathbf{R}^3$ open. Choose any parametrization $(u(t), v(t))$ for a *plane curve* contained in U. Then the composite

$$\alpha : I \to V \cap S, \ \alpha(t) = \mathbf{x}(u(t), v(t))$$

is, of course, a parametrization for a smooth curve contained in S. In fact, the opposite is also true, at least locally:

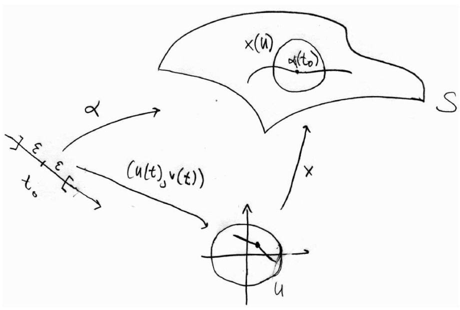


Figure 2.10: Curves on a surface

Proposition 2.33 *Let* $\alpha : I \to \mathbf{R}^3$ *denote a smooth curve on* S*. Let* $\mathbf{x} : U \to V \cap S$ *denote a parametrization with* $\mathbf{x}(u_0, v_0) = \alpha(t_0) \in S$*. Then there is an* $\varepsilon > 0$ *and a* smooth *parametrization* $(u(t), v(t))$ *for a curve in* U *defined on the interval* $(t_0 - \varepsilon, t_0 + \varepsilon) \subset \mathbf{R}$ *such that*

$$\alpha(t) = \mathbf{x}((u(t), v(t)) \text{ whenever } t_0 - \varepsilon < t < t_0 + \varepsilon. \tag{2.3}$$

Proof: We use Proposition 2.26 assuring existence of a local smooth "inverse" $F_0 : V_0 \to U_0$ to $\mathbf{x}$ defined on an open subset $V_0 \subset \mathbf{R}^3$ with $\alpha(t_0) \in V_0 \cap S$, $U_0 \subset U \subset \mathbf{R}^2$ open and $F_0 \circ \mathbf{x}(u, v) = (u, v)$ for all $(u, v) \in U_0$.

Choose $\varepsilon > 0$ such that $\alpha(t) \in V_0 \cap S$ for all $t_0 - \varepsilon < t < t_0 + \varepsilon$ (V_0 is open, α is continuous). The curve $\mathbf{x}^{-1}(\alpha(t)) = (u(t), v(t))$ is a continuous curve on U satisfying (2.3). We need to show that u and v are smooth functions. To this end, we note that

$$(u(t), v(t)) = \mathbf{x}^{-1}(\alpha(t)) = (F_0 \circ \mathbf{x})(\mathbf{x}^{-1}(\alpha(t))) = F_0(\alpha(t)))$$

is smooth as composition of the smooth maps F_0 and α for $t_0 - \varepsilon < t < t_0 + \varepsilon$.

□

Remark 2.34

- *Actually the curve $(u(t), v(t)) = \mathbf{x}^{-1} \circ \alpha(t)$ is smooth whenever it is defined. To prove this look at the above proof; we prove that $(u(t), v(t))$ is smooth in a neighborhood of all points t with $\alpha(t) \in \mathbf{x}(U)$; perhaps one has to use different local smooth "inverses" to the parametrization $\mathbf{x}$.*
- *Of course, one can define and investigate what it means that a curve is smooth at a particular point $\alpha(t_0)$ in the same manner. We will usually look at curves that are smooth on their entire domain.*

2.7 A useful trick: extending a chart to a local diffeomorphism

Still needed?

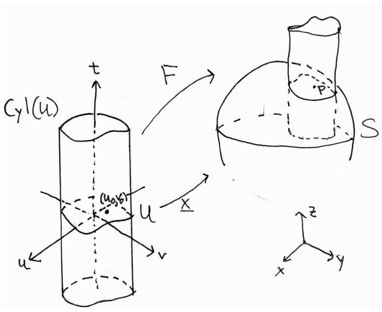


Figure 2.11: Illustration for Lemma 2.35

The cylinder over U and the image of this cylinder with a "cut", $F(u, v, t_0)$ to illustrate that $x(U)$ is just translated up and down.
Interactive illustration should allow the student to move this vertical cylinder around
on the sphere to see why we want the assumption on $D\mathbf{x}$.

Figure 2.12:

Lemma 2.35 *Let* $\mathbf{x} : U \to S$ *be a coordinate system around* $p \in S$. *Let* $p = \mathbf{x}(u_0, v_0)$ *and let the cylinder on* U *be* $Cyl(U) = \{(u, v, t) \in \mathbf{R}^3 | (u, v) \in U\}$. *Then there is a smooth map* $F : Cyl(U) \to \mathbf{R}^3$ *such that*

1. $F(u, v, 0) = \mathbf{x}(u, v)$
2. $dF_{(u_0,v_0,0)}$ *is a linear bijection.*
3. *There is an open set* $V \subseteq Cyl(U)$ *containing* $(u_0, v_0, 0)$ *and an open set* $W \subseteq \mathbf{R}^3$ *containing* p *such that* $F : V \to W$ *is a bijection with differentiable inverse.*

After perhaps choosing smaller V *and* W, *we have* $\mathbf{x}^{-1}_{|W \cap S} = \pi \circ F^{-1}_{|W \cap S}$, *where* $\pi(x, y, z) = (x, y)$

Proof: Once we have defined F to satisfy 1) and 2), 3) follows by the inverse function theorem. So what should F be? We can assume by renaming the axis if necessary, that the first two rows of $D\mathbf{x}_q$ are linearly independent. With this assumption, let $F(u, v, t) = (x_1(u, v), x_2(u, v), x_3(u, v)+t)$. The geometric content of this definition (cf. Fig. 2.12)

is that the cylinder over U is mapped to a cylinder over $\mathbf{x}(U)$. For fixed $t = t_0$, $F(u, v, t_0)$ is $\mathbf{x}(u, v)$ translated by the vector $(0, 0, t_0)$. The differential of F at $(u_0, v_0, 0)$ is

$$DF_{(u_0,v_0,0)} = \begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial u}(u_0, v_0) & \frac{\partial \mathbf{x}_1}{\partial v}(u_0, v_0) & 0 \\ \frac{\partial \mathbf{x}_2}{\partial u}(u_0, v_0) & \frac{\partial \mathbf{x}_2}{\partial v}(u_0, v_0) & 0 \\ \frac{\partial \mathbf{x}_3}{\partial u}(u_0, v_0) & \frac{\partial \mathbf{x}_3}{\partial v}(u_0, v_0) & 1 \end{pmatrix}$$

which has nonzero determinant. Hence F satisfies 1) and 2) and we conclude by the Inverse Function Theorem.

Now choose $\tilde{W}$ open such that $\tilde{W} \cap S = x(U)$; this is possible, since x is a coordinate chart. Let $\hat{W} = W \cap \tilde{W}$. Then on $\hat{W} \cap S$, $\mathbf{x}^{-1}$ is the restriction of the differentiable function $\pi \circ F^{-1}$.

$\square$

Notice that this is a sort of converse to Lemma 1.5 in the sense that the inverse of a local coordinate map, $\mathbf{x}^{-1}$, is in fact the restriction of $\Pi \circ F^{-1}$, at least in a small enough neighborhood of p.

3 Tangent planes and differentials

3.1 Linear approximation - definitions

A regular curve has a tangent line as its linear approximation at every point. Curvature measures the rate of change of the tangent directions. For a regular surface, we are now going to define tangent *planes*. Later on, we shall study rates of change for these planes (or rather their normals) that will provide us with measures for the curvature properties of a surface.

3.1.1 Tangent planes

Definition 3.1 *Let S denote a regular surface and $p \in S$ a point on S. The (linear)* tangent plane *T_pS to S at p consists of* all tangent vectors *to smooth curves* on *S* through *p (and at p), i.e.,*

$$T_pS = \{\alpha'(0) |\ \alpha :]-\varepsilon, \varepsilon[\to S,\ \varepsilon > 0,\ \alpha(0) = p\} \subset \mathbf{R}^3.$$

The affine *tangent plane Π_pS to S at p consists of all tangent* lines *to curves* on *S through p (and at p), i.e.,*

$$\Pi_\mathbf{p}S = p + T_\mathbf{p}S.$$

Remark 3.2 *You may think of the affine tangent plane Π_pS as the plane in 3-space that is best to approximate the surface S close to p. In analogy with the tangent line to a curve ¡t p that is best to approximate the curve close to p.*

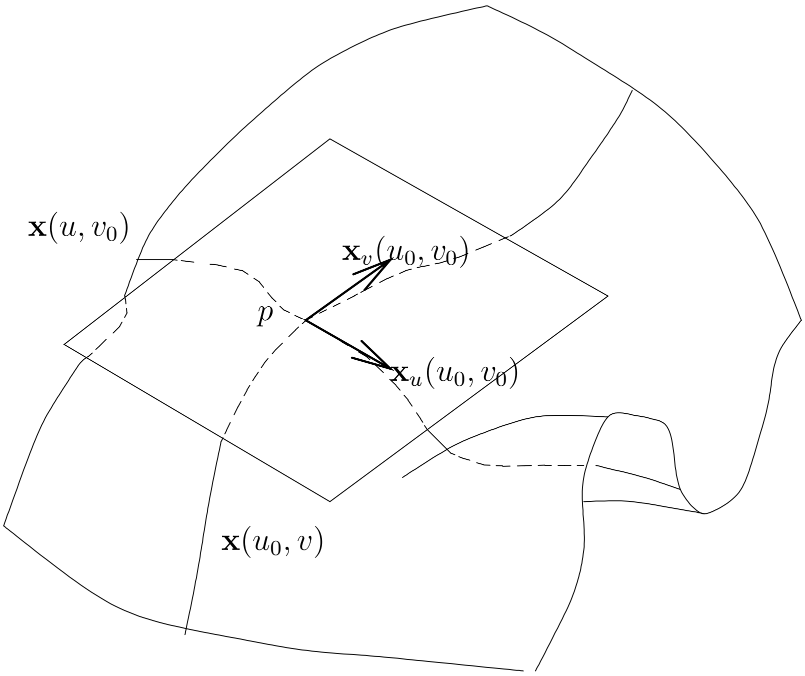


Figure 3.1: The tangent plane at p.

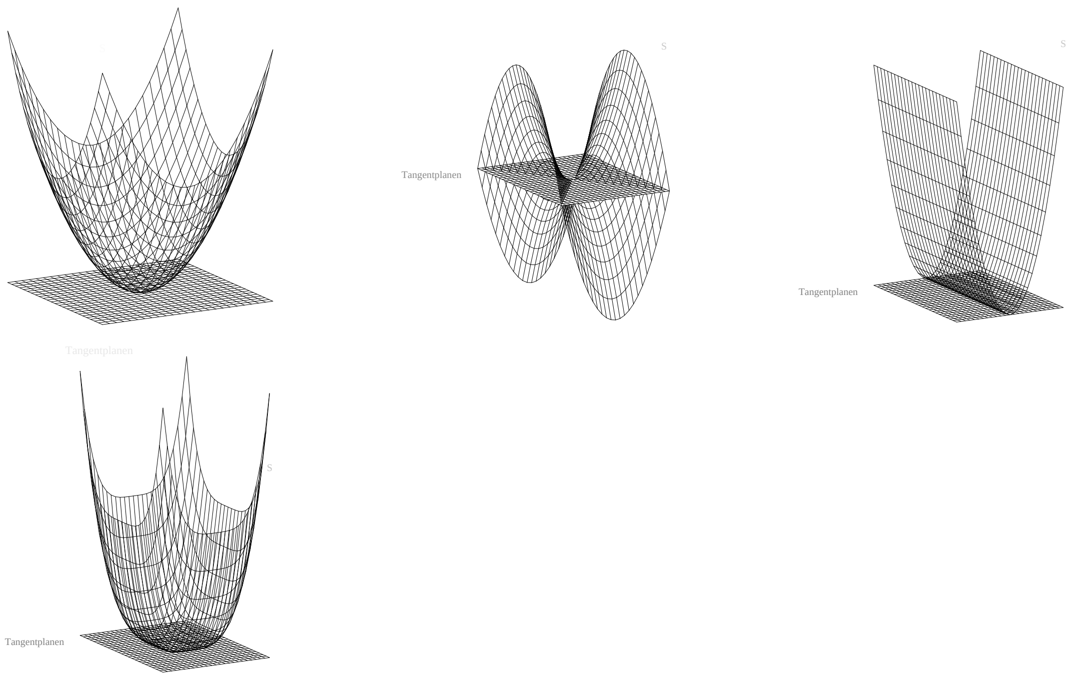


Figure 3.2: Tangent planes.

Our first aim is to show that both linear and affine tangent planes really are *planes* at every point of a regular surface:

Proposition 3.3 *Let S denote a regular surface and $p \in S$ a point on S. Then the (linear) tangent plane $T_pS \subset \mathbf{R}^3$ is a linear subspace of dimension 2. Moreover, if $\mathbf{x} : U \to V \cap S$ is a parametrization for S with $\mathbf{x}(q) = p$, the tangent plane agrees with the image of the Jacobian $d\mathbf{x}_q$:*

$$T_pS = d\mathbf{x}_q(\mathbf{R}^2) = sp(\mathbf{x}_u(q), \mathbf{x}_v(q)), \tag{3.1}$$

i.e., the plane spanned by the two linearly independent (!) vectors $\mathbf{x}_u(q), \mathbf{x}_v(q)$.

In other words, the tangent plane is the subspace *spanned by the tangent vectors to the parameter curves.*

Proof: It is enough to show (3.1): The image of the linear map $d\mathbf{x}_q$ is a linear subspace of $\mathbf{R}^3$. It has dimension two, since it is spanned by the two *linearly independent* vectors $\mathbf{x}_u(q)$ and $\mathbf{x}_v(q)$. We show (3.1) in two steps:

$T_pS \subseteq d\mathbf{x}_q(\mathbf{R}^2)$ Let $\mathbf{v} \in T_pS$. By definition, there is a parameterized curve $\alpha :]-\varepsilon, \varepsilon[\to S$ with $\alpha(0) = p$ and $\alpha'(0) = \mathbf{v}$. By Prop. 2.33, there is a plane curve with parametrization $(u(t), v(t))$, $t \in]-\varepsilon, \varepsilon[$[1] such that $q = (u(0), v(0))$ and $\alpha(t) = \mathbf{x}(u(t), v(t))$, $t \in]-\varepsilon, \varepsilon[$. The chain rule for differentiation yields

$$\begin{aligned} \mathbf{v} = \alpha'(0) &= \mathbf{x}_u(u(0), v(0))u'(0) + \mathbf{x}_v(u(0), v(0))v'(0) = \\ & \mathbf{x}_u(q)u'(0) + \mathbf{x}_v(q)v'(0) \in sp(\mathbf{x}_u(q), \mathbf{x}_v(q)). \end{aligned} \tag{3.2}$$

$d\mathbf{x}_q(\mathbf{R}^2) \subseteq T_pS$ Let $\mathbf{w} = a\mathbf{x}_u(q) + b\mathbf{x}_v(q)$, $a, b \in \mathbf{R}$. A glance at (3.2) shows that we can produce a curve $\alpha(t)$ on S through p with tangent vector $\mathbf{w}$ if we can find a curve $(u(t), v(t))$ in U with $q = (u(0), v(0))$ and $u'(0) = a$, $v'(0) = b$. This is easy: Choose $(u(t), v(t)) = q + (at, bt)$ in U on a sufficiently small interval $]-\varepsilon, \varepsilon[$.

□

Corollary 3.4 *With notation from above:*
$d\mathbf{x}_q : \mathbf{R}^2 \to T_pS$ is a linear isomorphism.

Proof: The vectorspace structure on the linear subspace T_pS is inherited from $\mathbf{R}^3$ and $d\mathbf{x}_q : \mathbf{R}^2 \to \mathbf{R}^3$ is a linear map, as it is the differential of a smooth map. Hence $d\mathbf{x}_q : \mathbf{R}^2 \to T_pS$ is a linear map. Moreover, $d\mathbf{x}_q$ is injective, since $\mathbf{x}$ is a parametrization, and it is surjective by 3.3.

[1] You might have to choose a smaller value of ε!

□

Remark 3.5 *Since* $d\mathbf{x}_q : \mathbf{R}^2 \to T_pS$ *is an isomorphism, it makes sense to talk about the inverse linear map* $(d\mathbf{x}_q)^{-1} : T_pS \to \mathbf{R}^2$ *and its values* $(d\mathbf{x}_q)^{-1}(w) \in \mathbf{R}^2$ *for* $w \in T_pS$*. Since* $\mathbf{x}^{-1}$ *is only defined as a continuous map on a subset of* S*, we cannot make sense of "*$d(\mathbf{x}^{-1})_p$*" however.*

For $w \in T_pS$*,* $w = \alpha'(0)$ *for a curve* $\alpha(t) = \mathbf{x}(u(t), v(t))$*, it follows from the proof above that* $(d\mathbf{x}_q)^{-1}(w) = (u'(0), v'(0))$*. Remark that the result does* not *depend on the particular choice of* α*; only on the tangent vector* $w = \alpha'(0)$*.*

3.1.2 The differential of a smooth function.

For a local analysis of smooth functions from $\mathbf{R}^k$ to $\mathbf{R}^n$ we use the differential from Definition 2.1. It was instrumental for instance statement and proof of the inverse and implicit function theorems and in general for approximations. For maps on surfaces, this local analysis will also be important. We define the differential of a smooth map from one regular surface to another. We leave it to the reader to spell an analogous definition for maps from regular surfaces to $\mathbf{R}^k$.

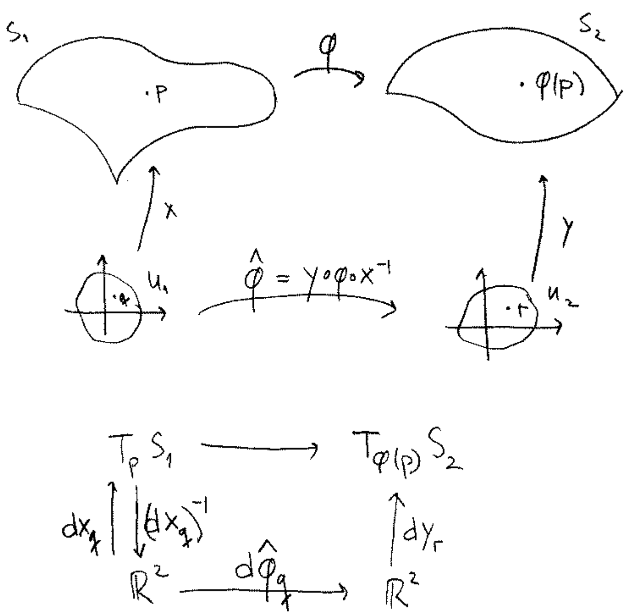


Figure 3.3: The differential of a map in local coordinates. The map $\hat{\varphi}$ stands for $\mathbf{y}^{-1} \circ \varphi \circ \mathbf{x}$

Definition 3.6 *Let $\varphi : S_1 \to S_2$ be a smooth map from the regular surface S_1 to the regular surface S_2 and let $p \in S_1$. Let $\alpha :]-\varepsilon, \varepsilon[\to S_1$ denote a smooth curve on S_1 with $\alpha(0) = p$. The differential $d\varphi_p$ is the map*

$$d\varphi_p : T_pS_1 \to T_{\varphi(p)}S_2$$

defined by $d\varphi_p(\alpha'(0)) = (\varphi \circ \alpha)'(0)$.

Proposition 3.7 *Definition 3.6 defines a linear map from T_pS_1 to $T_{\varphi(p)}S_2$.*

In other words,

- you represent a tangent vector $\mathbf{w}$ by a curve α on S_1 with $\alpha'(0) = \mathbf{w}$,
- you determine the image curve $\varphi \circ \alpha$ on S_2 under α, and
- you determine the tangent $(varphi \circ \alpha)'(0)$ of that image curve.

Proof: The proof consists of two steos:

1. The procedure above yields a well-defined map, i.e., given $\mathbf{w} \in T_pS_1$ with $\mathbf{w}w = \alpha'(0) = \tilde{\alpha}'(0)$ with α and $\tilde{\alpha}$ curves on S_1, then $(\varphi\circ\alpha)'(0) = (\varphi\circ\tilde{\alpha})'(0)$.
2. The map $d\varphi_p$ is linear.

The proof of both issues makes use of local coordinates: Let $\mathbf{x} : U_1 \to S_1$ be a coordinate system around p and let $\mathbf{y} : U_2 \to S_2$ be a coordinate system around $\varphi(p)$. According to Definition 2.30, U_1 can be chosen so that $\hat{\varphi} = \mathbf{y}^{-1} \circ \varphi \circ \mathbf{x} : U_1 \to U_2$ is a smooth map between open sets in the plane. Denote $\mathbf{x}^{-1}(p)$ by q and $\mathbf{y}^{-1}(\varphi(p))$ by r.

Let $\alpha(t) = \mathbf{x}(u(t), v(t))$ for t small enough, and define $\hat{\varphi}(u, v) = (\mathbf{y}^{-1} \circ \varphi \circ \mathbf{x})(u, v) = (\hat{\varphi}_1(u, v), \hat{\varphi}_2(u, v))$. We will prove that

$$d\varphi_p(\alpha'(0)) = d\mathbf{y}_r(d\hat{\varphi}_q((d\mathbf{x}_q)^{-1}(\alpha'(0)). \qquad (3.3)$$

By Cor. 3.4 and Rem. 3.5, the maps on the right hand side are all linear and independent of the choice of a curve α with the given tangent vector. This allows us to conclude 1. and 2. above.

To show (3.3), note that $d\varphi_p(\alpha'(0))\varphi \circ \mathbf{x} = \mathbf{y} \circ \hat{\varphi}$ and hence

$$d\varphi_p(\alpha'(0)) = \frac{d}{dt}(\varphi \circ \alpha)(0) = \frac{d}{dt}(\varphi \circ \mathbf{x}(u(t), v(t)))|_{t=0} = d\mathbf{y}_r(\frac{d}{dt}(\hat{\varphi}(u(t), v(t)))|_{t=0}$$

and

$$\frac{d}{dt}\hat{\varphi}(u(t),v(t))|_{t=0} = \frac{\partial\hat{\varphi}}{\partial u}(q)u'(0) + \frac{\partial\hat{\varphi}}{\partial v}(q)v'(0)$$
$$= \begin{pmatrix} \frac{\partial\hat{\varphi}_1}{\partial u}(q) & \frac{\partial\hat{\varphi}_1}{\partial v}(q) \\ \frac{\partial\hat{\varphi}_2}{\partial u}(q) & \frac{\partial\hat{\varphi}_2}{\partial v}(q) \end{pmatrix} \begin{pmatrix} u'(0) \\ v'(0) \end{pmatrix}$$

Hence $d\varphi_p(\alpha'(0)) = d\mathbf{y}_r(d\hat{\varphi}_q(u'(0),v'(0)) = d\mathbf{y}_r(d\hat{\varphi}_q(d\mathbf{x}_q)^{-1}(\alpha'(0))$.

□

3.1.3 Normal vectors

A vector $\mathbf{n} \in \mathbf{R}^3$ is called normal to S at the point p if $\mathbf{n}$ is perpendicular to every tangent vector $\mathbf{v} \in T_pS$. The set of all normal vectors at p is the 1-dimensional subspace $N_pS = T_pS^{\perp}$. Since N_pS is 1-dimensional, there are precisely *two unit* normal vectors to S at p - having opposite direction.

How can one calculate a (unit) normal vector? Let $x : U \to V \cap S$ denote a parametrization for S with $x(q) = p \in V \cap S$. The vectors $x_u(q)$ and $x_v(q)$ span the tangent plane T_pS; hence the vector $x_u(q) \times x_v(q)$ is perpendicular on both $x_u(q)$ and $x_v(q)$ and hence on every vector in T_pS; moreover, $x_u(q) \times x_v(q) \neq \mathbf{0}$ as the wedge product of linearly independent vectors. A *unit* normal vector is thus found as

$$N(p) = \frac{x_u(q) \times x_v(q)}{\|x_u(q) \times x_v(q)\|}.$$

If we choose another parametrization $y : U' \to V' \cap S$ with $y(q') = p$, then the vector $\frac{y_u(q') \times y_v(q')}{\|y_u(q') \times y_v(q')\|}$ is a unit normal vector, too. We conclude:

$$\frac{y_u(q') \times y_v(q')}{\|y_u(q') \times y_v(q')\|} = \pm\frac{x_u(q) \times x_v(q)}{\|x_u(q) \times x_v(q)\|}. \tag{3.4}$$

How can one find the correct sign in (3.4)?

Lemma 3.8 *The sign in (3.4) coincides with the sign of the determinant of the Jacobian of the change of coordinates map $h = y^{-1} \circ x$ at q.*[2]

Proof: By definition of the Jacobians and the chain rule, we obtain $d\mathbf{x}_q = d\mathbf{y}_{h(q)} \circ dh_q$ and hence the matrix equation

$$[x_u(q), x_v(q)] = [y_u(q'), y_v(q')] \circ Dh(q)$$

which by ex. 1 :implies: $x_u(q) \times x_v(q) = (y_u(q') \times y_v(q')) \det dh_q$.

□

[2] By the way: This sign is constant on (connected components of) the domain of h, since $detJh$ is a continuous function never attaining the value 0.

3.2 Oriented Surfaces

3.2.1 A Unit Normal Vector Field

Definition 3.9 *A unit normal vector field on a regular surface S is a smooth map $N : S \to \mathbf{R}^3$ such that for all $p \in S$*

- $|N(p)| = 1$ *and*
- *For all $v \in T_pS$, $N(p)$ is perpendicular to v.*

Not all surfaces have a normal vector field

Example 3.10 *The Möbius band.*

Remark 3.11 *If N is a unit normal vectorfield on S, then $-N$ is also a unit normal vectorfield, and these are the only two unit normal vectorfields on S, if S is connected.*

Definition 3.12 *A regular surface S is orientable if there is a unit normal vector field $N : S \to \mathbf{R}^3$ on S. A specific choice of a unit normal vector field on S is called an* orientation*. A surface with an orientation is an oriented surface.*

Theorem 3.13 *A regular surface S is orientable if and only if there is a set of parametrizations $\mathbf{x}^i : U_i \to S$, $i \in \mathcal{I}$ such that $\bigcup_{i \in \mathcal{I}} \mathbf{x}^i(U) = S$ and for any pair $i, j \in \mathcal{I}$ such that $\mathbf{x}^i(U_i) \cap \mathbf{x}^j(U_j) \neq \emptyset$, the Jacobian determinant $det(D((x^i)^{-1} \circ x^j)) > 0$*

Proof: Suppose S is orientable and let N be an orientation of S. Let $\mathbf{\tilde{x}^i} : \mathbf{U_i} \to \mathbf{S}$, $i \in \mathcal{I}$ be a set of parametrizations covering S and suppose each U_i is connected. Now construct the $\mathbf{x}^i : U_i \to S$ as follows: Let $p_i \in U_i$. If $N(p_i) = \frac{\tilde{x}^i_u \times \tilde{x}^i_v}{|\tilde{x}^i_u \times \tilde{x}^i_v|}$, then this holds for all $p_i \in U$ (U is connected) and we let $\mathbf{x}^i(u, v) = \tilde{x}^i(u, v)$, otherwise let $\mathbf{x}^i(u, v) = \tilde{x}^i(v, u)$.

The Jacobian for change of coordinates in this atlas is positive as the reader may see from exercise 1.

Suppose now that there is a set of parametrizations as stated in the theorem. Then define $N(p) = \frac{x^i_u \times x^i_v}{|x^i_u \times x^i_v|}$. This is well-defined by exercise 1, and it is clearly a unit normal vector field.

□

3.2.2 The Gauss map of an oriented manifold

3.2.3 Exercises

1. Let $\mathbf{v} = (v_1, v_2, v_3)$, $\mathbf{w} = (w_1, w_2, w_3)$, $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ and suppose that
$$\begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} \mathbf{X}$$
where $\mathbf{X}$ is a 2 by 2 matrix. Prove that $\mathbf{v} \times \mathbf{w} = \mathbf{a} \times \mathbf{b} \cdot \det(\mathbf{X})$

Proposition 3.14 *Let $S \subset \mathbf{R}^3$ denote a regular surface in 3-space, and let $p \in S$ denote an arbitrary point on S. Then there is an open subset $V \subseteq \mathbf{R}^3$ such that $V \cap S$ contains p and coincides with the image of a graph coordinate system.*

Proof: Let $\mathbf{x} : U \to \mathbf{R}^3$ denote an arbitrary parametrization of S with $U \subset \mathbf{R}^2$ open and $p \in \mathbf{x}(U)$. Let $q \in U$ such that $\mathbf{x}(q) = p$. Assume without loss of generality that *the first two* rows of the Jacobi-matrix $D\mathbf{x}_q$ are linearly independent.

Let $\pi : \mathbf{R}^3 \to \mathbf{R}^2$, $\pi(x, y, z) = (x, y)$ denote the linear map projecting 3-space vertically onto the XY-plane. We want to show, that there is an open subset $V_0 \subseteq \mathbf{R}^3$ containing p such that the restriction of π to $V_0 \cap S$ has a smooth inverse, which can serve as graph coordinate system.

Consider the composite map
$$\pi \circ \mathbf{x} : U \to \mathbf{R}^2, \quad (u, v) \mapsto (x_1(u, v), x_2(u, v)).$$
Its Jacobi-matrix is the 2-by-2 matrix consisting of the first two rows of $D\mathbf{x}$ (at every point in U). In particular, the differential $d(\pi \circ \mathbf{x})_q$ has non-zero determinant and hence it is a linear bijection. Applying the inverse function theorem to the map $\pi \circ \mathbf{x} : U \to \mathbf{R}^2$ yields:

There are open sets $U_0 \subset U$ containing q and $\tilde{U} \subset \mathbf{R}^2$ containing $\pi(p)$ such that the restriction $\pi \circ \mathbf{x} : U_0 \to \tilde{U}$ is a *diffeomorphism*. In particular, it has a smooth inverse $\psi : \tilde{U} \to U_0$. Our claim is, that the map
$$\mathbf{x} \circ \psi : \tilde{U} \to \mathbf{R}^3$$
is a graph coordinate system which certainly has p in its image. This is easiest explained by a look at the following commutative diagram:

$$[x_1(u, v), x_2(u, v), x_3(u, v)] \in \mathbf{x}(U_0)$$
$$\downarrow \pi$$
$$U_0 \ni (u, v) to^{\mathbf{x}} \xleftarrow[\psi]{} (x_1(u, v), x_2(u, v)) \in \tilde{U}$$

Since U_0 is open, by Def. 2.4.3 there is an open subset $V_0 \subseteq \mathbf{R}^3$ such that $\mathbf{x} : U_0 \to V_0 \cap S$ is a bijection and thus $\mathbf{x} \circ \psi : \tilde{U} \to V_0 \cap S$ is bijective. As a composite of smooth maps, the map $\mathbf{x} \circ \psi$ is smooth itself; moreover, as can be seen from the diagram, it preserves the first two coordinates, as required.

□

4 Metric on a surface: the first fundamental form

4.1 Distances, angles and areas on a surface

4.1.1 The metric on the tangent plane

4.1.2 The first fundamental form and linear algebra

Definition 4.1 *The 1. fundamental form I_p on T_pS is defined as the restriction of the usual dot-product in $\mathbf{R}^3$, i.e.,*

$$I_p : T_pS \times T_pS \to \mathbf{R}, \quad I_p(\mathbf{v}, \mathbf{w}) = \mathbf{v} \cdot \mathbf{w}.$$

Remark 4.2 *The associated quadratic form $\overline{I_p} : T_pS \to \mathbf{R}$ is given by $\overline{I_p}(\mathbf{v}) = \|\mathbf{v}\|^2$.*

Let us calculate Gram matrices (Def. 4.7) associated to I_p; those *depend on the choice of a basis* for T_pS. We assume given a parametrization $\mathbf{x} : U \to V \cap S$ with $\mathbf{x}(q) = p \in V$ giving rise to the basis $\{\mathbf{x}_u(q), \mathbf{x}_v(q)\}$. Since the dot-product in $\mathbf{R}^3$ is symmetric, so is the first fundamental form, and the coefficients of the

Gram matrix $B_{\mathbf{x}}$ are given as

$$\begin{aligned} E(p) &= I_p(\mathbf{x}_u(q), \mathbf{x}_u(q)) = \mathbf{x}_u(q) \cdot \mathbf{x}_u(q) \\ F(p) &= I_p(\mathbf{x}_u(q), \mathbf{x}_v(q)) = \mathbf{x}_u(q) \cdot \mathbf{x}_v(q) \\ G(p) &= I_p(\mathbf{x}_v(q), \mathbf{x}_v(q)) = \mathbf{x}_v(q) \cdot \mathbf{x}_v(q). \end{aligned}$$

In particular, we obtain for $\mathbf{v} = a_1\mathbf{x}_u(q) + a_2\mathbf{x}_v(q) \in T_pS$ and $\mathbf{w} = b_1\mathbf{x}_u(q) + b_2\mathbf{x}_v(q) \in T_pS$:

$$I_p(\mathbf{v}, \mathbf{w}) = [a_1, a_2] \begin{bmatrix} E(p) & F(p) \\ F(p) & G(p) \end{bmatrix} [b_1, b_2]^T.$$

Notice that

$$\begin{bmatrix} E(p) & F(p) \\ F(p) & G(p) \end{bmatrix} = (D\mathbf{x}_q)^T D\mathbf{x}_q$$

where $D\mathbf{x}_q = [\mathbf{x}_u(q), \mathbf{x}_v(q)]$ is the Jacobi matrix for $\mathbf{x}$ If we choose another parametrization $\mathbf{y} : U' \to V' \cap S$ with $\mathbf{y}(q') = p \in V'$, we have to determine the Gram matrix $B_{\mathbf{y}} = \begin{bmatrix} \bar{E}(p) & \bar{F}(p) \\ \bar{F}(p) & \bar{G}(p) \end{bmatrix}$ with respect to the basis $\{\mathbf{y}_u(q'), \mathbf{y}_v(q')\}$. The connection between the two Gram-matrices is given by

Corollary 4.3 $B_{\mathbf{x}} = D(\mathbf{y}^{-1} \circ \mathbf{x})_q^T B_{\mathbf{y}} D(\mathbf{y}^{-1} \circ \mathbf{x})_q$.

Proof:

$$B_{\mathbf{x}} = (D\mathbf{x}_q)^T D\mathbf{x}_q = (D(\mathbf{y} \circ \mathbf{y}^{-1} \circ \mathbf{x})_q)^T D(\mathbf{y} \circ \mathbf{y}^{-1} \circ \mathbf{x})_q =$$

$$(D\mathbf{y}_{q'} D(\mathbf{y}^{-1} \circ \mathbf{x})_q)^T D\mathbf{y}_{q'} D(\mathbf{y}^{-1} \circ \mathbf{x})_q = (D(\mathbf{y}^{-1} \circ \mathbf{x})_q)^T (D\mathbf{y}_{q'})^T D\mathbf{y}_{q'} D(\mathbf{y}^{-1} \circ \mathbf{x})_q =$$

$$D(\mathbf{y}^{-1} \circ \mathbf{x})_q^T B_{\mathbf{y}} D(\mathbf{y}^{-1} \circ \mathbf{x})_q.$$

□

4.1.3 The first fundamental form and arc length

4.1.4 The first fundamental form and area

4.2 Appendix on bilinear forms

4.2.1 Bilinear forms. Quadratic forms

Definition 4.4 *Let V denote a real vector space.*

1. *A map* $B : V \times V \to \mathbf{R}$ *is called a* bilinear form *if and only if it is linear in both variables, i.e., if*

   (a) $B(a\mathbf{v}_1 + b\mathbf{v}_2, \mathbf{w}) = aB(\mathbf{v}_1, \mathbf{w}) + bB(\mathbf{v}_2, \mathbf{w})$;

   (b) $B(\mathbf{v}, a\mathbf{w}_1 + b\mathbf{w}_2) = aB(\mathbf{v}, \mathbf{w}_1) + bB(\mathbf{v}, \mathbf{w}_2)$.

   for all vectors $\mathbf{v}, \mathbf{w}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1, \mathbf{w}_2 \in V$ *and real numbers* a, b.

2. *A bilinear form* $B : V \times V \to \mathbf{R}$ *is called*

   symmetric *if and only if* $B(\mathbf{v}, \mathbf{w}) = B(\mathbf{w}, \mathbf{v})$ *for all* $\mathbf{v}, \mathbf{w} \in V$.

   non-degenerate *if* $B(\mathbf{v}, \mathbf{w}) = 0$ *for all* $\mathbf{w} \in V$ *implies:* $\mathbf{v} = \mathbf{0}$.

   positive definite *if* $B(\mathbf{v}, \mathbf{v}) > 0$ *for all* $\mathbf{0} \neq \mathbf{v} \in V$.

3. *A bilinear form* $B : V \times V \to \mathbf{R}$ *induces a* quadratic form $\bar{B} : V \to \mathbf{R}$ *defined as* $\bar{B}(\mathbf{v}) = B(\mathbf{v}, \mathbf{v})$.

Example 4.5 1. *The usual dot-product on* $\mathbf{R}^n$ *is a positive definite symmetric bilinear form on* $\mathbf{R}^n$ *(and thus non-degenerate).*

2. *The map* $B' : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$, $B'([x_1, x_2], [y_1, y_2]) = x_1 y_1$ *is a degenerate symmetric bilinear form:* $B([0, 1], [y_1, y_2]) = 0$ *for all* $[y_1, y_2] \in \mathbf{R}^2$.

3. *The quadratic forms associated to the examples above are*

   (a) $\bar{B} : \mathbf{R}^n \to \mathbf{R}$, $\bar{B}(\mathbf{x}) = \|\mathbf{x}\|^2$.

   (b) $\bar{B}' : \mathbf{R}^2 \to \mathbf{R}$, $\bar{B}'([x_1, x_2]) = x_1^2$.

Remark 4.6 1. *A bilinear form* $B : V \times V \to \mathbf{R}$ *can be made symmetric using the definition* $B_s : V \times V \to \mathbf{R}$, $B_s(\mathbf{v}, \mathbf{w}) = \frac{1}{2}(B(\mathbf{v}, \mathbf{w}) + B(\mathbf{w}, \mathbf{v}))$. *The quadratic forms associated to* B *and to* B_s *agree:* $\bar{B} = \bar{B}_s$.

2. *The following formula shows how to recover a symmetric bilinear form from the quadratic form it induces:*

$$B(\mathbf{v}, \mathbf{w}) = \frac{1}{4}(\bar{B}(\mathbf{v} + \mathbf{w}) - \bar{B}(\mathbf{v} - \mathbf{w})).$$

4.2.2 Bilinear forms and Gram matrices

Let $\mathbf{E} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ denote a basis for the *finite-dimensional* vector space V. Let $B : V \times V \to \mathbf{R}$ denote a bilinear form on V.

Definition 4.7 *The Gram matrix* $\mathbf{B_E}$ *associated to* B *is the* $n \times n$*-matrix with entries* $(\mathbf{B_E})_{ij} = B(\mathbf{v}_i, \mathbf{v}_j)$.

Example 4.8 *The Gram-matrices associated to the bilinar forms in the examples above with respect to standard bases* $\mathbf{E}_n$ *on* $\mathbf{R}^n$*, resp.* $\mathbf{E}_2$ *on* $\mathbf{R}^2$ *are*

1. $\mathbf{B}_{\mathbf{E}_n} = \mathbf{I}_n$*, the identity matrix;*
2. $\mathbf{B}'_{\mathbf{E}_2} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

A linear map is determined by its values on all basis vectors. In the same spirit, a bilinear form is determined by the values on all *pairs* of basis vectors, i.e., by the associated Gram matrix. This is expressed in

Lemma 4.9 *Let* $B : V \times V \to \mathbf{R}$ *denote a bilinear form on the vector space* V*. Let* $\mathbf{E} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *denote a basis for* V*, and let* $\mathbf{B_E}$ *denote the associated Gram matrix with respect to* E*. Let* $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ *and* $\mathbf{w} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$ *be two vectors in* V*. Then*

$$B(\mathbf{v}, \mathbf{w}) = [a_1, \ldots, a_n]\mathbf{B_E}[b_1, \ldots, b_n]^T.$$

Proof: Straightforward calculation using the bilinearity of B.

□

4.2.3 Change of coordinates

Let V denote a vector space with the two bases $\mathbf{E} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\mathbf{F} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$. There is a uniquely determined regular $n \times n$-matrix $\mathbf{S}$ (change-of-base matrix) with the property: $[\mathbf{v}_1, \ldots, \mathbf{v}_n]\mathbf{S} = [\mathbf{w}_1, \ldots, \mathbf{w}_n]$. Coordinates wrt. the two bases are then transformed with the inverse matrix $\mathbf{S}^{-1}$ (change-of-coordinates matrix), i.e., if $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = a'_1\mathbf{w}_1 + \cdots + a'_n\mathbf{w}_n$, then

$$\mathbf{S}[a'_1, \ldots, a'_n]^T = [a_1, \ldots a_n]^T;$$

$$[a'_1, \ldots, a'_n]^T = \mathbf{S}^{-1}[a_1, \ldots, a_n]^T. \tag{4.1}$$

Let $B : V \times V \to \mathbf{R}$ denote a bilinear form on V. How are the Gram matrices $\mathbf{B_E}$ and $\mathbf{B_F}$ associated to B with respect to the two bases related to each other?

Proposition 4.10 $\mathbf{B_F} = \mathbf{S}^T\mathbf{B_E}\mathbf{S}$.

Proof: Let $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = a_1'\mathbf{w}_1 + \cdots + a_n'\mathbf{w}_n,\ \mathbf{w} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n = b_1'\mathbf{w}_1 + \cdots + b_n\mathbf{w}_n \in V$. We obtain:

$$[a_1, \ldots, a_n]\mathbf{B_E}[b_1, \ldots b_n]^T = B(\mathbf{v}, \mathbf{w}) = [a_1', \ldots, a_n']\mathbf{B_F}[b_1', \ldots b_n']^T.$$

Applying (4.1) to the left hand side, we obtain:

$$[a_1', \ldots, a_n']\mathbf{S}^T\mathbf{B_E}\mathbf{S}[b_1, \ldots b_n]^T = [a_1', \ldots, a_n']\mathbf{B_F}[b_1', \ldots b_n']^T.$$

Since this is correct for all column vectors $[a_1', \ldots, a_n']^T, [b_1', \ldots b_n']^T \in \mathbf{R}^n$, we conclude:

$$\mathbf{B_F} = \mathbf{S}^T\mathbf{B_E}\mathbf{S}.$$

□

4.2.4 Diagonalization

A matrix version of the spectral theorem from linear algebra can be phrased as follows:

Proposition 4.11 *Let* $\mathbf{A}$ *denote a real* symmetric $n \times n$*-matrix. Then there exists an orthonormal eigenvector basis* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *of* V*, i.e., there are real eigenvalues* λ_i *with*

$$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i.$$

Another way to phrase this property is in terms of the change of base matrix associated to the standard basis and the eigenvector basis:

Corollary 4.12 *There exist a diagonal matrix* $\boldsymbol{\Delta} = diag(\lambda_1, \ldots, \lambda_n)$ *and an othogonal matrix* $\mathbf{S} \in O(n)$ *such that*

$$A = \mathbf{S}^{-1}\boldsymbol{\Delta}\mathbf{S}.$$

Proof: Apply the change of base matrix $\mathbf{S}$ with $\mathbf{S}\mathbf{e}_i = \mathbf{v}_i$ to get: $\mathbf{SA} = \boldsymbol{\Delta}\mathbf{S}$.

□

There is an interesting application to quadratic forms (amongst others used in the classification of conic sections):

Corollary 4.13 *Let $\bar{B} : V \to \mathbf{R}$ denote the quadratic form associated to a bilinear form $B : V \times V \to \mathbf{R}$. There exists an orthormal basis $\mathbf{E} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of V and real numbers $\lambda_1, \ldots, \lambda_n$ such that*

$$\bar{B}(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = \lambda_1 a_1^2 + \cdots \lambda_n \mathbf{a}_n^2.$$

Proof: According to Cor. 4.12, the Gram matrix $\mathbf{B}_\mathbf{E}$ associated to B with respect to the eigenvector basis $\mathbf{E}$ is a diagonal matrix.

□

5 Curvature functions on regular surfaces

We already know how to measure the curvature of a curve, and we want to use this idea to measure the curvature of a surface. But there are some problems:

- Through each point of a surface there are infinitely many curves.
- Even if the surface is a plane, there will be for instance very small circles, i.e., curves with large curvature, and this is not really a result of "curvature" of the plane.

We will find out how to separate curvature of these curves into *normal curvature* which is curvature resulting from the surface being curved and *geodesic curvature*, which comes from the curve being curved relative to the surface.

Moreover, it turns out, that there is a systematic description of all the possible normal curvatures of curves through a given point on the surface. This will imply that the curvature properties of the infinitely many curves through a given point, can be captured by just two real numbers.

5.1 Vector fields

We will only need very special cases of vector fields here, but since the general concept is easy to define, we do that:

Definition 5.1 *A smooth vector field on a regular surface S is a smooth map $\chi : S \to \mathbf{R}^3$. A* tangent vector field *on S is a vectorfield such that $\chi(p) \in T_pS$ for all $p \in S$.*

Lemma 5.2 *Let $\chi : S \to \mathbf{R}^3$ be a tangent vector field on S and let $\mathbf{x} : U \to S$ be a local parametrization. Then $\chi \circ \mathbf{x} : U \to \mathbf{R}^3$ can be written $\chi \circ \mathbf{x}(u,v) = a(u,v)\mathbf{x}_u(u,v) + b(u,v)\mathbf{x}_y(u,v)$, where $a, b : U \to \mathbf{R}$ are smooth functions.*

Given smooth functions $a, b : U \to \mathbf{R}$, then $\chi(p) = a(\mathbf{x}^{-1}(u,v))\mathbf{x}_u(\mathbf{x}^{-1}(p)) + b(\mathbf{x}^{-1}(u,v))\mathbf{x}_v(\mathbf{x}^{-1}(p))$ is a smooth vector field on $\mathbf{x}(U) \subset S$

Proof: Since χ is smooth, $\chi \circ \mathbf{x}$ is smooth, by definition of smooth functions. Moreover, $\chi \circ \mathbf{x}(u,v) \in T_{\mathbf{x}(u,v)}S$, so $a(u,v)$ and $b(u,v)$ are uniquely given as the coordinates of $\chi \circ \mathbf{x}(u,v)$ in the basis $\mathbf{x}_u(u,v), \mathbf{x}_v(u,v)$. These are smooth, since

$$\begin{pmatrix} a(u,v) \\ b(u,v) \end{pmatrix} = \begin{pmatrix} E(u,v) & F(u,v) \\ F(u,v) & G(u,v) \end{pmatrix}^{-1} \begin{pmatrix} (\chi \circ \mathbf{x}(u,v)) \cdot \mathbf{x}_u(u,v) \\ (\chi \circ \mathbf{x}(u,v)) \cdot \mathbf{x}_v(u,v) \end{pmatrix}$$

Given a, b smooth, the function $\chi(p) = a(\mathbf{x}^{-1}(u,v))\mathbf{x}_u(\mathbf{x}^{-1}(p)) + b(\mathbf{x}^{-1}(u,v))\mathbf{x}_v(\mathbf{x}^{-1}(p))$ is a smooth vectorfield on $\mathbf{x}(U)$, since $\chi \circ \mathbf{x}$ is smooth.

□

Example 5.3 *normal vector field $\mathbf{N} : S \to \mathbf{R}^3$*

5.2 Normal sections and normal curvature

Let S denote an oriented regular surface and let $p \in S$. How can we get hold on curvature of the surface S close to the point p? A single number can't possibly give the answer: The surface might curve differently in different directions. You had better look at one curvature for every direction – which yields infinitely many of them. It will turn out later, that it is enough to know just *two* of them.

Let $\mathbf{N}(p)$ denote the unit normal vector at p corresponding to the chosen orientation.

Definition 5.4 *Let $\mathbf{v} \in T_pS$ denote a tangent vector of length 1: $\|\mathbf{v}\| = 1$. The* normal plane $\eta_{\mathbf{v}}$ *through p in direction $\mathbf{v}$ is the (affine) plane through p spanned by the tangent vector $\mathbf{v}$ and the normal vector $\mathbf{N}(p)$. The* normal section $\alpha_{\mathbf{v}}$ *at p in direction $\mathbf{v}$ is the curve that is obtained as the intersection of the normal plane $\eta_{\mathbf{v}}$ and the surface S.*

Remark 5.5 *One may apply the implicit function theorem to show that sufficiently close to p, every normal section $\alpha_{\mathbf{v}}$ in fact is a* regular curve *on S. In the following, we will even assume that $\alpha_{\mathbf{v}}$ is parametrized by arc length and that $\alpha_{\mathbf{v}}(0) = p$. Moreover, its (unit) tangent vector is contained in both T_pS and in the normal plane $\eta_{\mathbf{v}}$. Hence, by choosing the proper direction, we obtain: $\alpha'_{\mathbf{v}}(0) = \mathbf{v}$. What about the second derivative $\alpha''_{\mathbf{v}}(0)$? It is contained in the plane $sp(\mathbf{v}, \mathbf{N}(p))$ and perpendicular to $\mathbf{v}$ and hence has to be parallel to $\mathbf{N}(p)$.*

Example 5.6 *This is more or less an exercise for the reader: Let $S = \{(u, v, f(u, v))|(u, v) \in U\}$ and suppose that $f_u(u_0, v_0) = f_v(u_0, v_0) = 0$. Then $N(u_0, v_0) = (0, 0, 1)$ is a normal vector at (u_0, v_0) and all normal sections are graphs over a line $(u, v) = (u_0 + at, v_0 + bt)$. Hence a normal section is parametrized by $\alpha_{(a,b)}(t) = (u_0 + at, v_0 + bt, f(u_0 + at, v_0 + bt))$. These curves are regular and the normancurvatures can be calculated directly. Since all surfaces have local parametrizations as a graph, this example covers all cases.*
Exercise:

- *Prove that the normal sections are regular curves.*
- *Calculate the normal curvatures.*
- *Where is the implicit function theorem used here? (Remember, that it is equivalent to the inverse function theorem)*

Definition 5.7 *At the point p, the surface S has a* normal curvature *$k_n(\mathbf{v})$ in the tangent direction $\mathbf{v} \in T_pS$ given as the (signed) curvature of the normal section $\alpha_{\mathbf{v}}$:*

$$\alpha''_{\mathbf{v}}(0) = k_n(\mathbf{v})\mathbf{N}(p).$$

5.3 Normal curvature of arbitrary surface curves

While Def. 5.7 is best in order to understand the *meaning* of the term normal curvature, it is not nice to use it in calculations. In fact, the curves $\alpha_{\mathbf{v}}$ are only defined implicitly. To circumvent this problem, we define normal curvatures for *arbitrary* curves on the surface S: Let α denote a parametrization of a curve on S with unit speed parametrization, and assume $\alpha(0) = p$.

Definition 5.8 *The* normal curvature *of the curve α at p is defined as the (signed) length of the projection of the curvature vector $\alpha''_{\mathbf{v}}(0)$ on the normal vector $\mathbf{N}(p)$ of the surface:*

$$k_n(\alpha, p) = \alpha''_{\mathbf{v}}(0) \cdot \mathbf{N}(p).$$

As a dot product, this normal curvature can be calculated/interpreted using the curvature $\kappa(p)$ of the curve α at p and the *angle* θ between the principal normal vector $\mathbf{n}$ of the curve at P and the normal vector $\mathbf{N}(p)$ to the surface, to wit:

Lemma 5.9

$$k_n(\alpha, p) = \kappa(p) \cos\theta.$$

Remark 5.10 *For a* normal section*, the two definitions 5.7 and 5.8 agree:*

$$k_n(\alpha_{\mathbf{v}}, p) = k_n(\mathbf{v})\mathbf{N}(p) \cdot \mathbf{N}(p) = k_n(\mathbf{v}).$$

Proposition 5.11 *Normal curvature depends only on the tangent direction* $\mathbf{v}$ *and* not *on the curve* α*. In particular,*

$$k_n(\alpha, p) = k_n(\alpha'(0)) = II_p(\alpha'(0), \alpha'(0)).$$

Proof: The proof is contained in the section on moving tangent and normal vector fields, i.e., the calculation of normal curvature via the second fundamental form:

$$k_n(\alpha, p) = \alpha''(0) \cdot \mathbf{N}(p) = -dN_p(\alpha'(0)) \cdot \alpha'(0) = II_p(\alpha'(0), \alpha'(0)).$$

□

Outlook. The fact, that normal curvatures can be expressed via a bilinear form, is the key to using (bi)linear algebra towards their determination. The extremal *principal* curvatures are determined (up to sign) as the *eigenvalues* of the differential dN_p of the Gauss map; all other normal curvatures are *sandwiched* between the principal curvatures in a sense made precise by Euler's formula.

5.4 The Gauss map and its differential

5.4.1 Tangent and normal vector fields along a curve.

Some remarks aimed at giving a more systematic approach to the calculations of principal curvatures.

In this section, we study vector fields defined along a curve on a surface.

Definition 5.12 *Let S be a regular surface and let $\alpha : I \to S$ be a smooth curve on S. A smooth map: $\mathbf{w} : I \to \mathbf{R}^3$ is a tangent vector field along α if $\mathbf{w}(t) \in T_{\alpha(t)}S$ for all $t \in I$.*

Remark 5.13

- *When* $\mathbf{v} : S \to \mathbf{R}^3$ *is a tangent vector field, then* $\mathbf{v} \circ \alpha$ *is a tangent vector field along* α.
- *A vector field along a curve on* S *is a tangent vector field, if it consists of vectors which are tangent to the surface. An example is* $\mathbf{v}(t) = \alpha'(t)$. *This is then also tangent to the curve.*

Example 5.14 *Let* $\alpha(t) = (\cos(t), \sin(t), 0)$. *Then* α *is a curve on the surface* $S = \{(x, y, 0)|(x, y) \in \mathbf{R}^2\}$. *The vector field* $\mathbf{v}(t) = (1, 0, 0)$ *is a tangent vector field along* α *in* S, *since* $T_pS = \{(x, y, 0)|(x, y) \in \mathbf{R}^2\}$ *for all* $p \in S$.

Now consider α *as a curve in* S^2. *Then* $\mathbf{v}$ *is* not *a tangent vector field: For instance* $(1, 0, 0)$ *is not in* $T_{\alpha(0)}S^2 =$ *the* $x - z$ *plane.*

5.4.2 Normal curvature and the differential of the Gauss map

Lemma 5.15 *let* S *be an oriented regular surface with normal vector field* $\mathbf{N}$. *Let* $p \in S$ *and let* $\alpha : (-\varepsilon, \varepsilon) \to S$ *be a curve on* S *with* $\alpha(0) = p$. *Then* $\mathbf{N}(\alpha(t))$ *is the normal vector field along* α, *and we let* $\mathbf{w}(\alpha(t))$ *be a* tangent *vector field along the curve.*

Then

$$-dN_p(\alpha'(0)) \cdot \mathbf{w}(p) = \mathbf{N}(p) \cdot \frac{d}{dt}(\mathbf{w}(\alpha(t)))|_{t=0}.$$

Proof: The equation $\mathbf{N}(\alpha(t)) \cdot \mathbf{w}(\alpha(t)) = 0$ holds for all t. Take the derivative with respect to t.

□

Corollary 5.16 *Let* S *be an oriented regular surface and let* $v \in T_pS$. *Then*

$$II_p(\mathbf{v}, \mathbf{v}) = k_n(\mathbf{v}).$$

Proof: We insert in the lemma: Let $\alpha_{\mathbf{v}}$ be the arc length parametrization of a curve on S with $\alpha_{\mathbf{v}}(0) = p$ and $\alpha'_{\mathbf{v}}(0) = \mathbf{v}$; and let $\mathbf{w}(\alpha_{\mathbf{v}}(s)) = \alpha'_{\mathbf{v}}(s)$. When $\mathbf{v} \in T_pS$ we have

$$II_p(\mathbf{v}, \mathbf{v}) = -dN_p(\mathbf{v}) \cdot \mathbf{v} = -dN_p(c'_{\mathbf{v}}(0)) \cdot c'_{\mathbf{v}}(0) = \mathbf{N}(p) \cdot \frac{d}{ds} c'_{\mathbf{v}}|_{s=0} = \mathbf{N}(p) \cdot (\kappa(0)\mathbf{n}(0)) = k_n(\mathbf{v}).$$

Here $\mathbf{n}$ denotes the normal vector *of the curve*. When $\alpha_{\mathbf{v}}$ is a *normal section*, this is (plus or minus) the normal vector $N(p)$ of the surface.

□

Corollary 5.17 *The differential of the Gauss map, $dN_p : T_pS \to T_pS$ is self adjoint and the second fundamental form is symmetric. The coefficients in the hence symmetric Gram matrix for the second fundamental form are $e = N(p) \cdot \mathbf{x}_{uu}$, $f = N(p) \cdot \mathbf{x}_{vu}$ and $g = N(p) \cdot \mathbf{x}_{vv}$.*

Proof: Since $II_p(\mathbf{v}, \mathbf{w}) = -dN_p(\mathbf{v}) \cdot \mathbf{w} = -\mathbf{v} \cdot (dN_p)^*(\mathbf{w})$, where $(dN_p)^*$ is the adjoint of dN_p with respect to the inner product on $\mathbf{R}^3$, symmetry of II_p is equivalent to selfadjointness of dN_p. For this, it suffices to see, that $-dN_p(\mathbf{v}) \cdot \mathbf{w} = -\mathbf{v} \cdot (dN_p)(\mathbf{w})$ when $\mathbf{v}, \mathbf{w}$ are a basis for T_pS.

Insert the four combinations of the following (parameter) curves on S and tangent vectorfields along the curves in the lemma:

$$\alpha_u(t) = \mathbf{x}(u_0 + t, v_0),\ \alpha_v(t) = \mathbf{x}(u_0, v_0 + t),\ \mathbf{w} = \mathbf{x}_u,\ \mathbf{w} = \mathbf{x}_v.$$

This gives

1. e=$-dN_p(\mathbf{x}_u) \cdot (\mathbf{x}_u) = N(p) \cdot \mathbf{x}_{uu}$;
2. f=$-dN_p(\mathbf{x}_u) \cdot (\mathbf{x}_v) = N(p) \cdot \mathbf{x}_{vu}$;
3. f'=$-dN_p(\mathbf{x}_v) \cdot (\mathbf{x}_u) = N(p) \cdot \mathbf{x}_{uv} = f$;
4. g=$-dN_p(\mathbf{x}_v) \cdot (\mathbf{x}_v) = N(p) \cdot \mathbf{x}_{vv}$.

Notice that 2. og 3. give the same result; this proves that dN_p is self adjoint.

□

5.4.3 The second fundamental form

Let $N : S \to S^2$ denote the Gauss map for an oriented regular surface S and let $p \in S$.

Definition 5.18 *The second fundamental form on the tangent plane T_pS is given by*

$$II_p : T_pS \times T_pS \to \mathbf{R}, \quad II_p(\mathbf{v}, \mathbf{w}) = -I_p(dN_p(\mathbf{v}), \mathbf{w}) = -dN_p(\mathbf{v}) \cdot \mathbf{w}.$$

Remark 5.19 *The associated quadratic form $\overline{II_p} : T_pS \to \mathbf{R}$ is given by $\overline{II_p}(\mathbf{v}) = -dN_p(\mathbf{v}) \cdot \mathbf{v}$.*

Since dN_p is self adjoint, the second fundamental form is symmetric, and so is the Gram matrix with respect to any basis. The Gram matrix associated to II_p with respect to the basis $\{\mathbf{x}_u(q), \mathbf{x}_v(q)\}$ arising from a parametrization $\mathbf{x} : U \to V \cap S$ as above has coefficients

$$\begin{aligned} e(p) &= II_p(\mathbf{x}_u(q), \mathbf{x}_u(q)) = -dN_P(\mathbf{x}_u(q)) \cdot \mathbf{x}_u(q) \\ f(p) &= II_p(\mathbf{x}_u(q), \mathbf{x}_v(q)) = -dN_P(\mathbf{x}_u(q)) \cdot \mathbf{x}_v(q) \\ g(p) &= II_p(\mathbf{x}_v(q), \mathbf{x}_v(q)) = -dN_P(\mathbf{x}_v(q)) \cdot \mathbf{x}_v(q). \end{aligned}$$

In the following section, these coefficients will be calculated in more familiar terms. Moreover, we will uncover the connection between the quadratic form $\overline{II_p}$ and the normal curvatures at p.

5.4.4 An alternative treatment of normal and geodesic curvature via the Darboux frame

For an arc-length parametrized curve $\gamma : I \to S$ on a surface S, let $\mathbf{t}(s)$ denote the unit tangent vector field and $\mathbf{N}(s) = \mathbf{N}(\gamma(s))$ the normal vector along the curve. We complete it to yield a 3-dimensional frame field by defining $\hat{\mathbf{t}}(s) = \mathbf{N}(s) \times \mathbf{t}(s)$. As a result, we achive the so-called *Darboux frame*, a matrix-valued function $D : I \to SO(3)$ given by $D(s) = [\mathbf{t}(s), \hat{\mathbf{t}}(s), \mathbf{N}(s)]$; in particular, $D(s)D^T(s) = D^T(s)D(s) = I_3$ for all $s \in I$.

As for the Frenet frame, important information is hidden in the relation between the function $D(s)$ and its derivative $\dot{D}(s) = [\dot{\mathbf{t}}(s), \dot{\hat{\mathbf{t}}}(s), \dot{\mathbf{N}}(s)]$ in the form $\dot{D}(s) = D(s)B(s)$ with $B(s) = D^T(s)\dot{D}(s)$.

Lemma 5.20 *$B(s)$ is skew-symmetric for every $s \in I$.*

Proof: As for the Frenet frame, cf. Lemma 1.64.

□

We calculate $B(s) = D^T(s)\dot{D}(s)$ using the "row times column" formula for the matrix-matrix multiplication to yield:

$$B = \begin{bmatrix} \mathbf{t}\cdot\dot{\mathbf{t}} & \mathbf{t}\cdot\dot{\hat{\mathbf{t}}} & \mathbf{t}\cdot\dot{\mathbf{N}} \\ \hat{\mathbf{t}}\cdot\dot{\mathbf{t}} & \hat{\mathbf{t}}\cdot\dot{\hat{\mathbf{t}}} & \hat{\mathbf{t}}\cdot\dot{\mathbf{N}} \\ \mathbf{N}\cdot\dot{\mathbf{t}} & \mathbf{N}\cdot\dot{\hat{\mathbf{t}}} & \mathbf{N}\cdot\dot{\mathbf{N}} \end{bmatrix}$$

Now we exploit that the matrix B is skew-symmetric and that we know its first row in terms of the geodesic and the normal curvature of the curve γ (at a given value s): $\kappa_g = \dot{\mathbf{t}}\cdot\hat{\mathbf{t}}, \kappa_n = \dot{\mathbf{t}}\cdot\mathbf{N}$. Hence,

$$B = \begin{bmatrix} 0 & -\kappa_g & -\kappa_n \\ \kappa_g & 0 & -\tau_g \\ \kappa_n & \tau_g & 0 \end{bmatrix}$$

The third entity τ_g is called the *geodesic torsion* of the curve γ and will be identified later. First, we observe the new information about geodesic curvature in the $(1,3)$ entry of the matrix B. It tells us that $\kappa_n = -\mathbf{t}\cdot\dot{\mathbf{N}}$. What is $\dot{\mathbf{N}}$? In fact, it stands short for

$$\frac{d}{ds}(\mathbf{N}\circ\gamma) = D_\gamma\mathbf{N}\dot{\gamma} = D_\gamma\mathbf{N}\mathbf{t}$$

using the chain rule – with $\mathbf{N} : S \to S^2$ the Gauss map[1]. Using the convention for the Weingarten map $W = -D\mathbf{N}$, we arrive at the following expression of normal curvature κ_n in terms of the *second* fundamental form $<,>_{II}$.

Proposition 5.21 *The normal curvature κ_n of a regular surface S at a point $p \in S$ in (unit) tangent direction $\mathbf{t} \in T_pS$ has the form*

$$\kappa_n = -\mathbf{t}\cdot\dot{\mathbf{N}} = \mathbf{t}\cdot W(\mathbf{t}) = <\mathbf{t},\mathbf{t}>_{II}$$

Exercise. Calculate the geodesic torsion τ_g of an arc-length parametrized curve γ on S (at a point with non-zero curvature) by expressing it through the curve's torsion $\tau(s)$ and the derivative $\dot{\alpha}(s)$ of the angle function $\alpha(s)$ with $\alpha(s)$ denoting the variable angle between the principal normal vector $\mathbf{n}(s)$ (to the curve) and the surface normal $\mathbf{N}(s)$ along the curve.

Result. $\tau_g(s) = \tau(s) + \dot{\alpha}(s)$.

Hint. Express the relation between the Darboux matrix $D(s)$ and the Frenet matrix $F(s)$ in the form of a matrix equation $D(s) = F(s)C(s)$. The matrix $C(s)$ expresses a (variable) rotation by the angle $\alpha(s)$ in the normal plane $\mathbf{t}(s)^\perp$.

[1]sometimes denoted $\mathcal{G}$

5.5 Calculating the Weingarten matrix - linear algebra

This section gives a calculation of the matrix of the differential of the Gauss map, $dN_p : T_pS \to T_{N(p)}S^2 = T_pS$, with respect to the basis $\{\mathbf{x}_u, \mathbf{x}_v\}$ of the tangentplane T_pS. This 2×2-matrix is called the *Weingarten*-matrix $\mathbf{W}$.

Recall that the two fundamental forms I_p and II_p on the tangentplane are defined as follows: Let $\mathbf{v}, \mathbf{w} \in T_pS$, then

$$\begin{aligned} I_p(\mathbf{v}, \mathbf{w}) &= \mathbf{v} \cdot \mathbf{w} \\ II_p(\mathbf{v}, \mathbf{w}) &= -I_p(dN_p(\mathbf{v}), \mathbf{w}) = -dN_p(\mathbf{v}) \cdot \mathbf{w}. \end{aligned}$$

In the basis $\{\mathbf{x}_u, \mathbf{x}_v\}$ for the tangentplane T_pS we have a description via the Gram-matrices of these *symmetric* bilinear forms:

$$\begin{aligned} I_p(\mathbf{v}, \mathbf{w}) &= \mathbf{v}^T \begin{bmatrix} E & F \\ F & G \end{bmatrix} \mathbf{w}, \\ II_p(\mathbf{v}, \mathbf{w}) &= \mathbf{v}^T \begin{bmatrix} e & f \\ f & g \end{bmatrix} \mathbf{w}, \end{aligned}$$

where $\mathbf{v}$ and $\mathbf{w}$ are now the coordinates of these vectors in the basis $\{\mathbf{x}_u, \mathbf{x}_v\}$. Where is it used that the bilinear forms are symmetric?

The linear map $dN_p : T_pS \to T_pS$ is given in this basis by the matrix $\mathbf{W}$, i.e..: $dN_p(\mathbf{v}) = \mathbf{W}\mathbf{v}$ for $\mathbf{v} \in T_pS$.

By definition of the second fundamental form, we have

$$\mathbf{v}^T \begin{bmatrix} e & f \\ f & g \end{bmatrix} \mathbf{w} = II_p(\mathbf{v}, \mathbf{w}) = -I_p(dN_p(\mathbf{v}), \mathbf{w}) = -(\mathbf{W}\mathbf{v})^T \begin{bmatrix} E & F \\ F & G \end{bmatrix} \mathbf{w} = -\mathbf{v}^T\mathbf{W}^T \begin{bmatrix} E & F \\ F & G \end{bmatrix} \mathbf{w}.$$

Since this holds for *all* vectors $\mathbf{v}, \mathbf{w} \in T_pS$, we conclude:

$$\begin{bmatrix} e & f \\ f & g \end{bmatrix} = -\mathbf{W}^T \begin{bmatrix} E & F \\ F & G \end{bmatrix}.$$

Transpose this matrix equation and multiply by the inverse of the Gram-matrix for the 1. fundamentalform. This gives

$$\begin{aligned} \mathbf{W} &= -\begin{bmatrix} E & F \\ F & G \end{bmatrix}^{-1} \begin{bmatrix} e & f \\ f & g \end{bmatrix} \\ &= \frac{1}{EG - F^2} \begin{bmatrix} -G & F \\ F & -E \end{bmatrix} \begin{bmatrix} e & f \\ f & g \end{bmatrix}. \end{aligned}$$

The Gauss curvature, $K(p)$, is the determinant of the linear map dN_p, so:

$$K(p) = \det \mathbf{W} = \frac{\det \begin{bmatrix} e & f \\ f & g \end{bmatrix}}{\det \begin{bmatrix} E & F \\ F & G \end{bmatrix}} = \frac{eg - f^2}{EG - F^2}.$$

And we can calculate the *mean curvature*, which is defined by $H(p) = -\frac{\text{tr } dN_p}{2} = -\frac{\text{tr } \mathbf{w}}{2}$, so:

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

Using these two curvatures, we can now calculate the *principal curvatures* k_1 og k_2 at the point p. Since k_1 og k_2 are the eigenvalues of the linear map $-dN_p$ they are roots of the characteristic polynomium $x^2 + (\text{tr } dN_p)x + \det dN_p = x^2 - 2H(p)x + K(p)$,i.e.:

$$k_{1/2}(p) = H(p) \pm \sqrt{H(p)^2 - K(p)}.$$

5.6 Gauss curvature and mean curvature

5.6.1 Classification of points on a surface

5.6.2 An outlook at minimal surfaces

6 Isometries. Invariance of the Gauss curvature

6.1 Isometries between surfaces

6.2 Theorema egregium and consequences

6.3 Christoffel symbols

This is a section on a technical device which is indispensable both in the proof of Gauss' Theorema egregium and when handling geodesics and geodesic curvature.

6.3.1 Definition

In the following, we fix a parametrization $\mathbf{x} : U \to V \cap S$ for the surface S. At every point $p = \mathbf{x}(q)$, the three vectors $\mathbf{x}_u(q), \mathbf{x}_v(q)$ and $\mathbf{N}(p)$ form a (moving) basis for $\mathbf{R}^3$. Let us express the double derivatives $x_{uu}(q), x_{uv}(q)$ and $x_{vv}(q)$ as linear combinations of these basis vectors (we omit the ps and qs and regard the derivatives etc. as vector fields):

$$\begin{aligned} \mathbf{x}_{uu} &= \Gamma_{11}^1\mathbf{x}_u + \Gamma_{11}^2\mathbf{x}_v + \Gamma_{11}^3\mathbf{N} \\ \mathbf{x}_{uv} &= \Gamma_{12}^1\mathbf{x}_u + \Gamma_{12}^2\mathbf{x}_v + \Gamma_{12}^3\mathbf{N} \\ \mathbf{x}_{vv} &= \Gamma_{22}^1\mathbf{x}_u + \Gamma_{22}^2\mathbf{x}_v + \Gamma_{22}^3\mathbf{N}. \end{aligned} \tag{6.1}$$

with unknown real coefficient functions $\Gamma_{ij}^k : U \to \mathbf{R}$. Since the vector fields $\mathbf{x}_u, \mathbf{x}_v$ and N are linearly independent, the system (6.1) has a *unique* solution at every point $p \in S$. The equations 6.1 can be written in matrix form as

$$[\mathbf{x}_{uu}\mathbf{x}_{uv}\mathbf{x}_{vv}] = [\mathbf{x}_u\mathbf{x}_v\mathbf{N}] \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \\ \Gamma_{11}^3 & \Gamma_{12}^3 & \Gamma_{22}^3 \end{bmatrix} \tag{6.2}$$

6.3.2 The metric determines the Christoffel symbols

We start by differentiating the equations $E = \mathbf{x}_u \cdot \mathbf{x}_u,\ F = \mathbf{x}_u \cdot \mathbf{x}_v,\ G = \mathbf{x}_v \cdot \mathbf{x}_v$ with respect to u and v to get:

$$E_u = 2\mathbf{x}_u \cdot \mathbf{x}_{uu} \quad F_u = \mathbf{x}_u \cdot \mathbf{x}_{uv} + \mathbf{x}_v \cdot \mathbf{x}_{uu} \quad G_u = 2\mathbf{x}_v \cdot \mathbf{x}_{uv} \tag{6.3}$$

$$E_v = 2\mathbf{x}_u \cdot \mathbf{x}_{uv} \quad F_v = \mathbf{x}_u \cdot \mathbf{x}_{vv} + \mathbf{x}_v\mathbf{x}_{uv} \quad G_v = 2\mathbf{x}_v \cdot \mathbf{x}_{vv}. \tag{6.4}$$

Solving with respect to the six dot products involved yields the matrix equation
$\begin{bmatrix} \frac{E_u}{2} & \frac{E_v}{2} & F_v - \frac{G_u}{2} \\ F_u - \frac{E_v}{2} & \frac{G_u}{2} & \frac{G_v}{2} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_u \cdot \mathbf{x}_{uu} & \mathbf{x}_u \cdot \mathbf{x}_{uv} & \mathbf{x}_u \cdot \mathbf{x}_{vv} \\ \mathbf{x}_v \cdot \mathbf{x}_{uu} & \mathbf{x}_v \cdot \mathbf{x}_{uv} & \mathbf{x}_v \cdot \mathbf{x}_{vv} \end{bmatrix} = [\mathbf{x}_u\mathbf{x}_v]^T[\mathbf{x}_{uu}\mathbf{x}_{uv}\mathbf{x}_{vv}] =$
$[\mathbf{x}_u\mathbf{x}_v]^T[\mathbf{x}_u\mathbf{x}_v\mathbf{N}] \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \\ \Gamma_{11}^3 & \Gamma_{12}^3 & \Gamma_{22}^3 \end{bmatrix} = \begin{bmatrix} E & F & 0 \\ F & G & 0 \end{bmatrix} \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \\ \Gamma_{11}^3 & \Gamma_{12}^3 & \Gamma_{22}^3 \end{bmatrix} = \begin{bmatrix} E & F \\ F & G \end{bmatrix} \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \end{bmatrix}$
and hence:

$$\begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \end{bmatrix} = \begin{bmatrix} E & F \\ F & G \end{bmatrix}^{-1} \begin{bmatrix} \frac{E_u}{2} & \frac{E_v}{2} & F_v - \frac{G_u}{2} \\ F_u - \frac{E_v}{2} & \frac{G_u}{2} & \frac{G_v}{2} \end{bmatrix}.$$

In particular, these six Christoffel symbols can be expressed as (rational) functions in the coefficients of the first fundamental form and their partial derivatives of the first order. In particular:

Proposition 6.1 *The Christoffel-symbols* Γ_{ij}^k, $1 \le i, j, k \le \mathbf{2}$ *are intrinsic entities, i.e., they depend only on the metric (first fundamental form) along the surface* S.

To determine the last three Christoffel symbols Γ^3_{ij} observe that they are given by the coefficients $e, f, g$of the second fundamental form (and hence *not* intrinsic):
$[e\ f\ g] = [\mathbf{N}\cdot\mathbf{x}_{uu}\ \mathbf{N}\cdot\mathbf{x}_{uv}\ \mathbf{N}\cdot\mathbf{x}_{vv}] = \mathbf{N}^T[\mathbf{x}_u\mathbf{x}_v\mathbf{N}]\begin{bmatrix}\Gamma^1_{11} & \Gamma^1_{12} & \Gamma^1_{22}\\ \Gamma^2_{11} & \Gamma^2_{12} & \Gamma^2_{22}\\ \Gamma^3_{11} & \Gamma^3_{12} & \Gamma^3_{22}\end{bmatrix} = [001]\begin{bmatrix}\Gamma^1_{11} & \Gamma^1_{12} & \Gamma^1_{22}\\ \Gamma^2_{11} & \Gamma^2_{12} & \Gamma^2_{22}\\ \Gamma^3_{11} & \Gamma^3_{12} & \Gamma^3_{22}\end{bmatrix} =$
$[\Gamma^3_{11}\ \Gamma^3_{12}\ \Gamma^3_{22}]$.

6.4 A proof of Theorema egregium

6.4.1 Proof of the Gauss equations and the Mainardi-Codazzi equations

OBS: In this section, we use L, M, N instead of e, f, g as notation for the coefficients of the 2nd fundamental form with respect to a local coordinate system σ.

Start with the equation $\mathbf{0} = (\sigma_{uu})_v - (\sigma_{uv})_u$. Expressing the double derivatives in terms of Christoffel symbols and differentiating results in:
$\mathbf{0} = (\Gamma^1_{11}\sigma_u + \Gamma^2_{11}\sigma_v + L\mathbf{N})_v - (\Gamma^1_{12}\sigma_u + \Gamma^2_{12}\sigma_v + M\mathbf{N})_u =$
$= ((\Gamma^1_{11})_v - (\Gamma^1_{12})_u)\sigma_u + ((\Gamma^2_{11})_v - (\Gamma^2_{12})_u)\sigma_v + (L_v - M_u)\mathbf{N} - \Gamma^1_{12}\sigma_{uu} + (\Gamma^1_{11} - \Gamma^2_{12})\sigma_{uv} +$
$\Gamma^2_{11}\sigma_{vv} + L\mathbf{N}_v - M\mathbf{N}_u =$
$= [\sigma_u\ \sigma_v\ \mathbf{N}](\begin{bmatrix}(\Gamma^1_{11})_v - (\Gamma^1_{12})_u\\ (\Gamma^2_{11})_v - (\Gamma^2_{12})_u\\ L_v - M_u\end{bmatrix} + \begin{bmatrix}\Gamma^1_{11} & \Gamma^1_{12} & \Gamma^1_{22}\\ \Gamma^2_{11} & \Gamma^2_{12} & \Gamma^2_{22}\\ L & M & N\end{bmatrix}\begin{bmatrix}-\Gamma^1_{12}\\ \Gamma^1_{11} - \Gamma^2_{12}\\ \Gamma^2_{11}\end{bmatrix} - \begin{bmatrix}a & b\\ c & d\\ 0 & 0\end{bmatrix}\begin{bmatrix}-M\\ L\end{bmatrix})$.
The last term uses the Weingarten matrix W:
$[\mathbf{N}_u\ \mathbf{N}_v] = [\sigma_u\ \sigma_v]\begin{bmatrix}a & b\\ c & d\end{bmatrix} = -[\sigma_u\ \sigma_v]W.$
The equation corresponding to the $\mathbf{N}$-coordinate yields:
$0 = L_v - M_u + \begin{bmatrix}L\\ M\\ N\end{bmatrix}\cdot\begin{bmatrix}-\Gamma^1_{12}\\ \Gamma^1_{11} - \Gamma^2_{12}\\ \Gamma^2_{11}\end{bmatrix}.$
The two equations corresponding to the σ_u, σ_v-coordinates yield:
$W\begin{bmatrix}-M\\ L\end{bmatrix} = \begin{bmatrix}(\Gamma^1_{11})_v - (\Gamma^1_{12})_u\\ (\Gamma^2_{11})_v - (\Gamma^2_{12})_u\end{bmatrix} + \begin{bmatrix}\Gamma^1_{11} & \Gamma^1_{12} & \Gamma^1_{22}\\ \Gamma^2_{11} & \Gamma^2_{12} & \Gamma^2_{22}\end{bmatrix}\begin{bmatrix}-\Gamma^1_{12}\\ \Gamma^1_{11} - \Gamma^2_{12}\\ \Gamma^2_{11}\end{bmatrix}.$
The left hand side is equal to
$W\begin{bmatrix}-M\\ L\end{bmatrix} = \mathcal{F}_1^{-1}\mathcal{F}_2\begin{bmatrix}-M\\ L\end{bmatrix} = \mathcal{F}_1^{-1}\begin{bmatrix}L & M\\ M & N\end{bmatrix}\begin{bmatrix}-M\\ L\end{bmatrix} = \frac{1}{EG-F^2}\begin{bmatrix}G & -F\\ -F & E\end{bmatrix}\begin{bmatrix}0\\ LN - M^2\end{bmatrix} =$
$= \frac{LN-M^2}{EG-F^2}\begin{bmatrix}-F\\ E\end{bmatrix} = \begin{bmatrix}-FK\\ EK\end{bmatrix}.$

7 Geodesics and the Gauss-Bonnet theorem

7.1 Geodesic curvature. Geodesics

7.1.1 Definitions

7.1.2 The geodesic differential equations

A curve γ on a surface S is geodesic if and only if the curve's curvature vectors $\ddot{\gamma}(t)$ are parallel to the surface normals $\mathbf{N}(\gamma(t))$ for all $t \in I$, i.e., each $\ddot{\gamma}(t)$ is perpendicular on the tangent space $T_{\gamma(t)}S$. Using a coordinate patch $\mathbf{x} : U \to S$ we have (possibly after restriction):

1. $\gamma(t) = \mathbf{x}(u(t), v(t))$ for a curve $(u(t), v(t))$ in U
2. γ geodesic $\Leftrightarrow \ddot{\gamma} \cdot (\mathbf{x}_u \circ \gamma) = 0 = \ddot{\gamma} \cdot (\mathbf{x}_v \circ \gamma)$ (for all $s \in I$).

Calculations

Applying the chain rule and the product rule yields:

1. $\dot{\gamma} = \dot{u}\mathbf{x}_u + \dot{v}\mathbf{x}_v$

2. $\ddot{\gamma} = \ddot{u}\mathbf{x}_u + \ddot{v}\mathbf{x}_v + \dot{u}^2\mathbf{x}_{uu} + 2\dot{u}\dot{v}\mathbf{x}_{uv} + \dot{v}^2\mathbf{x}_{vv}$

We use the (3×3)-matrix function Γ with Christoffel symbol functions(!) $\Gamma_{ij}^k,\ 1 \leq i \leq j \leq 2,\ 1 \leq k \leq 3$, as coefficients (cf. 6.2), ie. the matrix equation

$$[\mathbf{x}_{uu}\mathbf{x}_{uv}\mathbf{x}_{vv}] = [\mathbf{x}_u\mathbf{x}_v\mathbf{N}] \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \\ \Gamma_{11}^3 & \Gamma_{12}^3 & \Gamma_{22}^3 \end{bmatrix}$$

using the 3D-coordinate systems $\{\mathbf{x}_u, \mathbf{x}_v, \mathbf{N}\}$ at every point in U, resp. $\mathbf{x}(U)$. The two first matrices above have the given vectors as *column* vectors. We continue the calculation from 2. above:

$$\ddot{\gamma} = [\mathbf{x}_u\mathbf{x}_v\mathbf{N}] \begin{bmatrix} \ddot{u} \\ \ddot{v} \\ 0 \end{bmatrix} + [\mathbf{x}_{uu}\mathbf{x}_{uv}\mathbf{x}_{vv}] \begin{bmatrix} \dot{u}^2 \\ 2\dot{u}\dot{v} \\ \dot{v}^2 \end{bmatrix} = [\mathbf{x}_u\mathbf{x}_v\mathbf{N}](\begin{bmatrix} \ddot{u} \\ \ddot{v} \\ 0 \end{bmatrix} + \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \\ \Gamma_{11}^3 & \Gamma_{12}^3 & \Gamma_{22}^3 \end{bmatrix} \begin{bmatrix} \dot{u}^2 \\ 2\dot{u}\dot{v} \\ \dot{v}^2 \end{bmatrix}).$$

The geodesic condition (2) from the introduction translates to:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = [\mathbf{x}_u\mathbf{x}_v]^T\ddot{\gamma} = [\mathbf{x}_u\mathbf{x}_v]^T[\mathbf{x}_u\mathbf{x}_v\mathbf{N}](\begin{bmatrix} \ddot{u} \\ \ddot{v} \\ 0 \end{bmatrix} + \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \\ \Gamma_{11}^3 & \Gamma_{12}^3 & \Gamma_{22}^3 \end{bmatrix} \begin{bmatrix} \dot{u}^2 \\ 2\dot{u}\dot{v} \\ \dot{v}^2 \end{bmatrix}) =$$

$$= \begin{bmatrix} E & F & 0 \\ F & G & 0 \end{bmatrix} (\begin{bmatrix} \ddot{u} \\ \ddot{v} \\ 0 \end{bmatrix} + \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \\ \Gamma_{11}^3 & \Gamma_{12}^3 & \Gamma_{22}^3 \end{bmatrix} \begin{bmatrix} \dot{u}^2 \\ 2\dot{u}\dot{v} \\ \dot{v}^2 \end{bmatrix}) = \begin{bmatrix} E & F \\ F & G \end{bmatrix} (\begin{bmatrix} \ddot{u} \\ \ddot{v} \end{bmatrix} + \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \end{bmatrix} \begin{bmatrix} \dot{u}^2 \\ 2\dot{u}\dot{v} \\ \dot{v}^2 \end{bmatrix}).$$

Since the the matrix $\mathcal{F}_I = \begin{bmatrix} E & F \\ F & G \end{bmatrix}$ is invertible, this condition is equivalent to:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \ddot{u} \\ \ddot{v} \end{bmatrix} + \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \end{bmatrix} \begin{bmatrix} \dot{u}^2 \\ 2\dot{u}\dot{v} \\ \dot{v}^2 \end{bmatrix}.$$

Each of the two rows denotes one of the two (2nd order homogeneous) *geodesic differential equations* for the unknown functions $(u(t), v(t))$.

7.1.3 Existence and uniqueness

7.2 The local Gauss-Bonnet theorem

7.2.1 Relating geodesic curvature and Gaussian curvature

7.2.2 Green's Theorem

7.2.3 Hopf's Umlaufsatz

7.2.4 Proof: Putting it altogether

7.3 A note on geodesic triangles.

As an example for an interpretation of the local Gauss Bonnet theorem, we have immediately from McCleary 12.4

Corollary 7.1 *For a geodesic triangle $T \subset S$ contained in a coordinate system, we have*

$$\int\int_T K = \psi_0 + \psi_1 + \psi_2 - \pi$$

where ψ_i are the interior angles in the triangle.

Written in local coordinates, where $\mathbf{x} : U \to S$ is a parametrization containing T, the left hand side of the equation says

$$\int\int_{\mathbf{x}^{-1}(T)} K(u,v)|\mathbf{x}_u \times \mathbf{x}_v| dudv =$$

$$\int\int_{\mathbf{x}^{-1}(T)} K(u,v)\sqrt{EG - F^2} dudv$$

Now the area Area(T) is

$$\int\int_{\mathbf{x}^{-1}(T)} 1|\mathbf{x}_u \times \mathbf{x}_v| dudv = \int\int_{\mathbf{x}^{-1}(T)} \sqrt{EG - F^2} dudv$$

Let $\{T_n\}_{n \in \mathbb{N}}$ be a sequence of geodesic triangles converging to p, i.e., for any $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that $n \geq N$ implies $T_n \subset B(p, \varepsilon)$.

Then

$$\lim_{n\to\infty} \frac{\iint_{\mathbf{x}^{-1}(T)} K(u,v)\sqrt{EG-F^2}dudv}{\iint_{\mathbf{x}^{-1}(T)} \sqrt{EG-F^2}dudv} = K(p)$$

and hence

$$K(p) = \lim_{n\to\infty} \frac{1}{\text{Area}(T_n)}(\psi_{0n} + \psi_{1n} + \psi_{2n} - \pi)$$

This is clearly invariant under isometries, since geodesic curves, angles and areas are. Hence it provides another argument for Theorema Egregium.

7.4 The global Gauss-Bonnet theorem

7.4.1 The Euler characteristic of a surface

7.4.2 Proof of the global Gauss-Bonnet theorem