Elementary Differential Geometry: Curves and Surfaces

Edition 2008

Martin Raussen

Department of Mathematical Sciences, Aalborg University

Fredrik Bajersvej 7G, DK – 9220 Aalborg Øst, Denmark, +45 96 35 88 55

e-mail: raussen@math.aau.dk

First of all, I would like to thank my colleague Lisbeth Fajstrup for many discussion about these notes and for many of the drawings in this text.

The VIDIGEO-project has provided interactive and dynamical software for illustrations of curves and associated objects which are used in the first chapters of these notes; the interactive versions are accessible at `www.math.aau.dk/~raussen/VIDIGEO/GEOLAB` – or by clicking on the figures in the electronic version of this text.

The project was supported by Dansk Naturvidenskabscenter; the underlying Java-software was programmed by Robert Sinclair, PhD.

An example of how you can use the geometric laboratory is given in Sect. 2.5. A prototype for graphical software with illustrations of material related to surfaces is currently developed by Martin Qvist, Aalborg University.

Contents

CHAPTER 1

Plane and Space: Linear Algebra and Geometry

The purpose of this course is the study of *curves* and *surfaces*, and those are, in general, *curved*. Nevertheless, our main tools to understand and analyze these curved objects are (tangent) *lines* and *planes* and the way those change along a curve, resp. surface. This is why we start with a brief chapter assembling prerequisites from linear geometry and algebra. Most or all of these will be known to the reader from elementary courses and textbooks. We focus on geometric aspects of methods borrowed from linear algebra; proofs will only be included for those properties that are important for the future development.

1. Vectors and Products

1.1. Vectors. Our models for plane and space use the *Euclidean vector spaces* $\mathbf{R}^2$, resp. $\mathbf{R}^3$ with coordinate systems $\{\mathbf{i}, \mathbf{j}\}$, resp. $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Hence, a plane vector $\mathbf{x} \in \mathbf{R}^2$ has a unique representation

$$\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j}, \text{ in short } \mathbf{x} = [x_1, x_2], \tag{1.1}$$

while a space vector $\mathbf{y} \in \mathbf{R}^3$ has a unique representation

$$\mathbf{y} = y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}, \text{ in short } \mathbf{y} = [y_1, y_2, y_3]. \tag{1.2}$$

where $x_1, x_2, y_1, y_2, x_3, y_3$ are real numbers. A special vector is the *zero vector*

$$\mathbf{0} = [0, 0] \in \mathbf{R}^2, \text{ resp. } \mathbf{0} = [0, 0, 0] \in \mathbf{R}^3.$$

With respect to the given coordinate system, the sum of two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{R}^3$ is calculated componentwise:

$$\mathbf{x} + \mathbf{y} = [x_1, x_2, x_3] + [y_1, y_2, y_3] = [x_1 + y_1,\ x_2 + y_2,\ x_3 + y_3].$$

Similarly for the difference of two vectors.

A vector $\mathbf{x} \in \mathbf{R}^3$ may be multiplied with a real number (or scalar) a:

$$a\mathbf{x} = [ax_1, ax_2, ax_3], \quad a \in \mathbf{R}.$$

Together, these operations give rise to the following concepts:

DEFINITION 1.1. (1) A vector $\mathbf{w} = a\mathbf{x} + b\mathbf{y},\ a, b \in \mathbf{R}$ is called a *linear combination* of the vectors $\mathbf{x}$ and $\mathbf{y}$. A vector $\mathbf{w} = a\mathbf{x} + b\mathbf{y} + c\mathbf{z},\ a, b, c \in \mathbf{R}$ is called a *linear combination* of the vectors $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$.

(2) A linear combination $\mathbf{w} = a\mathbf{x} + b\mathbf{y} + c\mathbf{z}$ is called *non-trivial* if and only if at least one of the coefficients is not 0 : $a \neq 0$ or $b \neq 0$ or $c \neq 0$.

(3) The set of *all* linear combinations of a set of vectors is called their *span*:
- $sp(\mathbf{x}) = \{a\mathbf{x} |\ a \in \mathbf{R}\}$;
- $sp(\mathbf{x}, \mathbf{y}) = \{a\mathbf{x} + b\mathbf{y} |\ a, b \in \mathbf{R}\}$;
- $sp(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \{a\mathbf{x} + b\mathbf{y} + c\mathbf{z} |\ a, b, c \in \mathbf{R}\}$.

DEFINITION 1.2. A set of vectors in Euclidean plane or space is called *linearly dependent* if the zero-vector $\mathbf{0}$ is expressible as a non-trivial linear combination (Def. 1.1(2)) of the vectors in the set, and *linearly independent* else.

EXAMPLE 1.3.

(1) The set $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\} \subset \mathbf{R}^3$ is linearly independent since:

$$[0, 0, 0] = \mathbf{0} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k} = [a, b, c] \Rightarrow a = b = c = 0.$$

(2) The set $\{[1, 1, 1], [1, -1, 2], [2, 4, 1]\} \subset \mathbf{R}^3$ is linearly dependent, since

$$\mathbf{0} = 3[1, 1, 1] + (-1)[1, -1, 2] + (-1)[2, 4, 1]$$

is a non-trivial linear combination yielding $\mathbf{0}$.

More systematically, we have:

LEMMA 1.4. (1) *Two non-zero vectors* $\mathbf{x}, \mathbf{y}$ *are linearly dependent if and only if they are* parallel, *i.e , if there exists a number* $d \in \mathbf{R}$ *such that* $\mathbf{y} = d\mathbf{x}$. *In that case,* $sp(\mathbf{x}, \mathbf{y}) = sp(\mathbf{x})$.

(2) *Three non-zero vectors* $\mathbf{x}, \mathbf{y}, \mathbf{z}$ *are linearly dependent if and only if they are* coplanar, *i.e., if* $\mathbf{x}$ *and* $\mathbf{y}$ *are parallel or if there exist numbers* $d, e \in \mathbf{R}$ *such that* $\mathbf{z} = d\mathbf{x} + e\mathbf{y}$. *In that case,* $sp(\mathbf{x}, \mathbf{y}, \mathbf{z})$ *is equal to the span of* 1 *or* 2 *of these vectors.*

PROOF:

(1) If $\mathbf{0} = a\mathbf{x} + b\mathbf{y}$ and $b \neq 0$, then $\mathbf{y} = \frac{-a}{b}\mathbf{x}$.
(2) Let $\mathbf{0} = a\mathbf{x} + b\mathbf{y} + c\mathbf{z}$. If $c = 0$, then $\mathbf{x}$ and $\mathbf{y}$ are parallel. If $c \neq 0$, then $\mathbf{z} = \frac{-a}{c}\mathbf{x} + \frac{-b}{c}\mathbf{y}$.

□

Vectors are useful in the description of the *Euclidean plane* $\mathbf{E}^2$ and of *Euclidean space* $\mathbf{E}^3$. The most elementary objects in plane, resp. space, are its *points*. A connection between points in $\mathbf{E}^i$ and vectors in $\mathbf{R}^i$ is established as follows: Choose a distinguished point O as *origin* of the coordinate system. To any point P, we associate the vector $\overrightarrow{OP}$, and its coordinates; we write

$$P[x_1, x_2, x_3] \quad \text{if} \quad \overrightarrow{OP} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}. \tag{1.3}$$

Given two points P, Q in plane or space with

$$\overrightarrow{OP} = [x_1, x_2, x_3], \quad \overrightarrow{OQ} = [y_1, y_2, y_3].$$

Then, the vector $\overrightarrow{PQ}$ is given as the *difference vector*

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = [y_1, y_2, y_3] - [x_1, x_2, x_3] = [y_1 - x_1, y_2 - x_2, y_3 - x_3].$$

Its geometric interpretation is an *arrow* (directed line segment) starting at P and ending at Q. The definitions above have the consequence, that

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{OQ} - \overrightarrow{OP} + \overrightarrow{OR} - \overrightarrow{OQ} = \overrightarrow{PR}.$$

Geometrically, $\overrightarrow{PR}$ corresponds to the arrow from P to R, which is the diagonal in the *parallelogram* spanned by the arrows from P to Q, resp. Q to R in Fig. 1.

Conversely, let $P \in \mathbf{E}^i$ denote an arbitrary point. Then, any vector $\mathbf{x} = [x_1, x_2, x_3]$ may be interpreted as an arrow *with initial point* P: If $\overrightarrow{OP} = [y_1, y_2, y_3]$, define Q by $\overrightarrow{OQ} = \overrightarrow{OP} + \mathbf{x} = [x_1 + y_1, x_2 + y_2, x_3 + y_3]$. In fact, $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \mathbf{x}$.

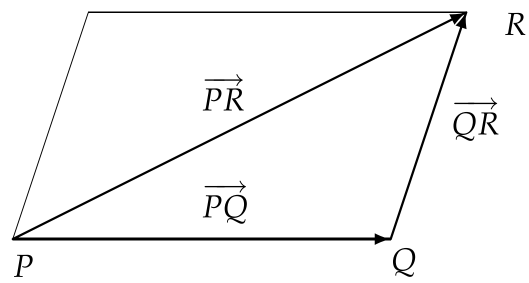


FIGURE 1. Addition of arrows

1.2. Products.

1.2.1. *The dot product.* The geometric concepts *length* of a vector and *angle* between two vectors are encoded in the *dot product* between two vectors: The dot product of two vectors $\mathbf{x} = [x_1, x_2, x_3]$ and $\mathbf{y} = [y_1, y_2, y_3]$ is given as the *real number*

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 \in \mathbf{R}. \tag{1.4}$$

The *length* of the vector $\mathbf{x}$ is defined as the non-negative real number

$$|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + x_3^2}. \tag{1.5}$$

Note that $|\mathbf{x}| = 0$ if and only if $\mathbf{x} = \mathbf{0}$. Of course, the definition of length relies on *Pythagoras* theorem. Its generalization, the *Law of Cosines*, is the background for the following geometric interpretation of the dot product of two vectors:

PROPOSITION 1.5. *Let α denote the angle between the two vectors $\mathbf{x}$ and $\mathbf{y}$. Then*

$$\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}|\,|\mathbf{y}| \cos \alpha. \tag{1.6}$$

The angle between two vectors in $\mathbf{R}^3$ has to be interpreted as the angle (between 0 and π or 180^o) in the plane that they span; it is zero, if they are parallel. Since the restriction of the real function cos to the interval $[0, \pi]$ – corresponding to angles between 0^o and 180^o – attains every value between -1 and 1 exactly once, Formula (1.6) can be used to recover the angle α between two non-zero vectors $\mathbf{x}$ and $\mathbf{y}$ from their dot product:

$$\cos \alpha = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|} = \frac{x_1 y_1 + x_2 y_2 + x_3 y_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}\sqrt{y_1^2 + y_2^2 + y_3^2}}. \tag{1.7}$$

The angle α between the two lines given by $\mathbf{x}$ and $\mathbf{y}$ is calculated as the arccos of that number – and is either a number in the interval $[0, \pi]$ or an angle between 0^0 and 180^0. Note that two vectors $\mathbf{x}$ and $\mathbf{y}$ are *perpendicular* to each other if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.

EXAMPLE 1.6.

(1) The coordinate vectors $\mathbf{i} = [1, 0, 0]$, $\mathbf{j} = [0, 1, 0]$, and $\mathbf{k} = [0, 0, 1]$ all have length 1 and are mutually perpendicular.
(2) Let $\mathbf{x} = [1, 2]$ and $\mathbf{y} = [-2, 4]$. Then, $|\mathbf{x}| = \sqrt{5}$; $|\mathbf{y}| = \sqrt{20} = 2\sqrt{5}$; and $\mathbf{x} \cdot \mathbf{y} = 6$. Hence, by Formula (1.7), the angle α between $\mathbf{x}$ and $\mathbf{y}$ satisfies: $\cos\alpha = \frac{6}{10} = \frac{3}{5}$, and hence $\alpha = \arccos\frac{3}{5}$, corresponding to an angle of 53.13^o.

1.2.2. *The plane product.* The *plane product* of two *plane* vectors $\mathbf{x} = [x_1, x_2]$, $\mathbf{y} = [y_1, y_2] \in \mathbf{R}^2$ is given by the determinant

$$[\mathbf{x}, \mathbf{y}] = x_1 y_2 - x_2 y_1 = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}. \tag{1.8}$$

This number can be interpreted as the (signed) area of the parallelogram spanned by $\mathbf{x}$ and $\mathbf{y}$, to wit: Let α denote the angle between the lines through $\mathbf{x}$ and $\mathbf{y}$. Then,

$$[\mathbf{x}, \mathbf{y}] = |\mathbf{x}||\mathbf{y}| \sin\alpha.$$

A *negative* plane product indicates thus, that the "shortest" rotation sending the half-line through $\mathbf{x}$ into the half-line through $\mathbf{y}$ is *clock-wise*.

The proof is similar to that of the corresponding formula for the wedge product (cf. (1.12)), but more elementary; it is therefore omitted here.

To a *plane* vector $\mathbf{x} = [x_1, x_2]$, we associate its *hat* vector $\hat{\mathbf{x}} = [-x_2, x_1]$. The reader should check that $\mathbf{x}$ and $\hat{\mathbf{x}}$ have the same length and are perpendicular to each other – using the dot product ((1.4) in Sect. 1.2.1). Only the plane vector $-\hat{\mathbf{x}} = [x_2, -x_1]$ has those same properties. They can be distinguished using the plane product from above: $[\mathbf{x}, \hat{\mathbf{x}}] = x_1^2 + x_2^2 \geq 0$, whereas $[\mathbf{x}, -\hat{\mathbf{x}}] \leq 0$. Geometrically, $\hat{\mathbf{x}}$ arises from $\mathbf{x}$ by a counter-clockwise rotation by an angle $\frac{\pi}{2}$ or 90^0.

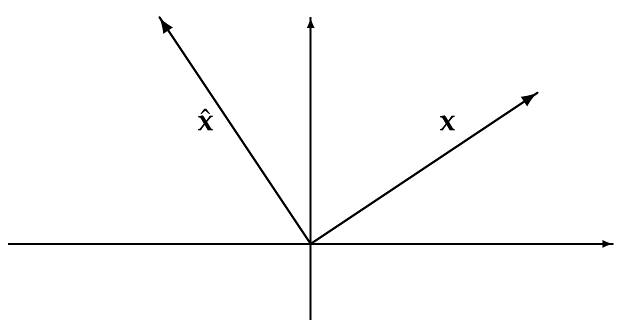


FIGURE 2. The hat vector

1.2.3. *The wedge product.* In space, we define the *cross product* (or *wedge product*) between two *space* vectors $\mathbf{x} = [x_1, x_2, x_3]$ and $\mathbf{y} = [y_1, y_2, y_3]$. This is a new *space vector* $\mathbf{x} \times \mathbf{y}$, whose coordinates are given by the following rule:

$$\mathbf{x} \times \mathbf{y} = [x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1]. \tag{1.9}$$

An easy way to memorize this definition is as the *determinant* of the following formal matrix with $\mathbf{i}, \mathbf{j}, \mathbf{k}$ denoting the coordinate vectors:

$$\mathbf{x} \times \mathbf{y} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} = (x_2y_3 - x_3y_2)\mathbf{i} - (x_1y_3 - x_3y_1)\mathbf{j} + (x_1y_2 - x_2y_1)\mathbf{k}. \tag{1.10}$$

1.2.4. *The space product.* The formal properties and a geometric interpretation of the wedge product are easier to derive using yet another product; a sort of strange product between *three* space vectors $\mathbf{z}$, $\mathbf{x}$ and $\mathbf{y}$: it is a real number, called the *scalar triple product* or *space product* between those three vectors:

$$[\mathbf{z}, \mathbf{x}, \mathbf{y}] = \mathbf{z} \cdot (\mathbf{x} \times \mathbf{y}) =$$
$$= z_1(x_2y_3 - x_3y_2) + z_2(x_3y_1 - x_1y_3) + z_3(x_1y_2 - x_2y_1) =$$

$$= \begin{vmatrix} z_1 & z_2 & z_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}. \tag{1.11}$$

1.2.5. *Properties of wedge and space product.* The vector product $\mathbf{x} \times \mathbf{y}$ is *perpendicular* to both $\mathbf{x}$ and $\mathbf{y}$, as follows from the calculations:

$$\mathbf{x} \cdot (\mathbf{x} \times \mathbf{y}) = \begin{vmatrix} x_1 & x_2 & x_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = 0,$$

$$\mathbf{y} \cdot (\mathbf{x} \times \mathbf{y}) = \begin{vmatrix} y_1 & y_2 & y_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = 0.$$

Why? The determinants above are zero, since they contain two *identical* rows.

REMARK 1.7.
These orthogonality properties can be used to check whether you made a mistake in your calculation of $\mathbf{x} \times \mathbf{y}$. Dot your result with both $\mathbf{x}$ and $\mathbf{y}$ – this is easy and not time consuming. If those dot products are not 0 – both of them – you certainly made an error. Try again!

The *length* of the wedge product $\mathbf{x} \times \mathbf{y}$ can be established from the following formula – which may be verified by a lengthy but routine verification using coordinates:

$$|\mathbf{x}|^2|\mathbf{y}|^2 = |\mathbf{x} \times \mathbf{y}|^2 + (\mathbf{x} \cdot \mathbf{y})^2. \tag{1.12}$$

Formulas (1.12) and (1.6) have the following consequence: Let α denote the angle between $\mathbf{x}$ and $\mathbf{y}$. Then

$$|\mathbf{x}\times\mathbf{y}|^2 = |\mathbf{x}|^2|\mathbf{y}|^2 - (\mathbf{x}\cdot\mathbf{y})^2 = |\mathbf{x}|^2|\mathbf{y}|^2 - |\mathbf{x}|^2|\mathbf{y}|^2\cos^2\alpha = \\ = |\mathbf{x}|^2|\mathbf{y}|^2(1-\cos^2\alpha) = |\mathbf{x}|^2|\mathbf{y}|^2\sin^2\alpha,$$

and hence

$$|\mathbf{x}\times\mathbf{y}| = |\mathbf{x}||\mathbf{y}|\sin\alpha. \tag{1.13}$$

To summarize, $\mathbf{x}\times\mathbf{y}$ is a vector that is perpendicular to both $\mathbf{x}$ and $\mathbf{y}$ and has lenght given by (1.13). The only other vector with these two properties is the vector $-(\mathbf{x}\times\mathbf{y})$. Those two can be distinguished by space products:
$[(\mathbf{x}\times\mathbf{y}),\mathbf{x},\mathbf{y}] = (\mathbf{x}\times\mathbf{y})\cdot(\mathbf{x}\times\mathbf{y}) = |\mathbf{x}\times\mathbf{y}|^2 \geq 0$, whereas $[-(\mathbf{x}\times\mathbf{y}),\mathbf{x},\mathbf{y}] = -|\mathbf{x}\times\mathbf{y}|^2 \leq 0$. Geometrically, this distinction is done by the *rule of thumbs*: $\mathbf{x},\mathbf{y}$ and $\mathbf{x}\times\mathbf{y}$ form a right-handed triple (use the first three fingers of your right hand to point in the direction of these vectors!)

The length of the cross product $\mathbf{x}\times\mathbf{y}$ has the following *geometric interpretation*: $|\mathbf{x}\times\mathbf{y}|$ is the *area* of the parallelogram (cf. Fig. 3) spanned by $\mathbf{x}$ and $\mathbf{y}$:

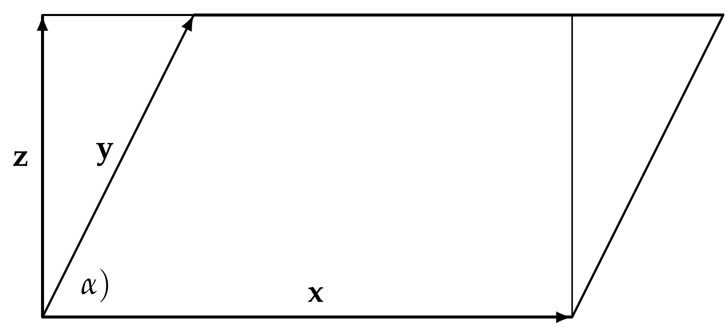


FIGURE 3. Cross product and area

Why? The vector $\mathbf{z}$ in Fig. 3 is chosen in the plane spanned by $\mathbf{x}$ and $\mathbf{y}$ and perpendicular to $\mathbf{x}$ with length $|\mathbf{z}| = |\mathbf{y}|\sin\alpha$. The parallelogram spanned by $\mathbf{x}$ and $\mathbf{y}$ has the same area A as the *rectangle* spanned by $\mathbf{x}$ and $\mathbf{z}$ (why?), which is
$A = |\mathbf{x}||\mathbf{z}| = |\mathbf{x}||\mathbf{y}|\sin\alpha = |\mathbf{x}\times\mathbf{y}|$.

The space product itself is interpreted as a (signed) *volume*, cf. Fig. 4: Let $\mathbf{z},\mathbf{x}$, and $\mathbf{y}$ denote vectors with the same initial point. Unless they are linearly dependent (coplanar) (cf. Lemma 1.4.2), they span a parallelepiped. The area of the base parallelogram spanned by $\mathbf{x}$ and $\mathbf{y}$ is given by $A = |\mathbf{x}\times\mathbf{y}|$. Now let α be the angle between $\mathbf{z}$ and $\mathbf{x}\times\mathbf{y}$. Assume for a moment that α is acute. Let $\mathbf{w}$ be the indicated (height) vector on the line through $\mathbf{x}\times\mathbf{y}$. Its length is given by $|\mathbf{z}|\cos\alpha$. The parallelepiped spanned by $\mathbf{x},\mathbf{y}$, and $\mathbf{z}$, resp. the one spanned by $\mathbf{x},\mathbf{y}$, and $\mathbf{w}$ have the same volume V, and thus:

$$V = |\mathbf{w}|A = |\mathbf{z}|\cos\alpha A = |\mathbf{z}||\mathbf{x}\times\mathbf{y}|\cos\alpha = \mathbf{z}\cdot(\mathbf{x}\times\mathbf{y}) = [\mathbf{z},\mathbf{x},\mathbf{y}].$$

If α happens not to be acute, one may replace α by $\theta = \pi - \alpha$. Since $\cos\theta = -\cos\alpha$, the corresponding calculation has the result: $V = -[\mathbf{z}, \mathbf{x}, \mathbf{y}]$.

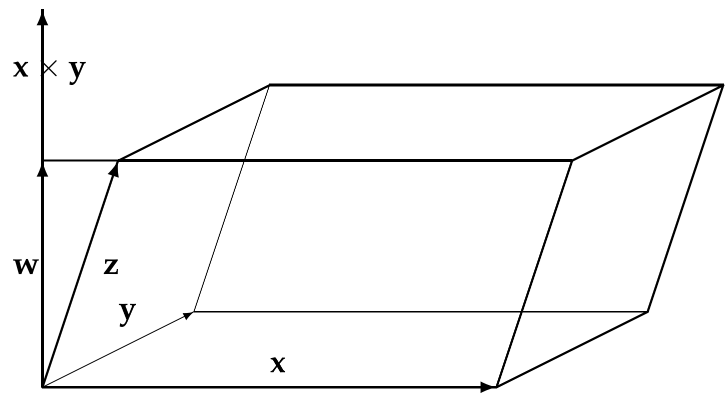


FIGURE 4. Space product and parallelepiped

EXAMPLE 1.8.
Three points in space are given by $P = [1, 2, 4]$, $Q = [-2, 3, -5]$ and $R = [0, 1, -1]$. Then $\overrightarrow{PQ} = [-2-1, 3-2, -5-4] = [-3, 1, -9]$, $\overrightarrow{PR} = [0-1, 1-2, -1-4] = [-1, -1, -5]$, and

$$\overrightarrow{PQ} \times \overrightarrow{PR} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -3 & 1 & -9 \\ -1 & -1 & -5 \end{vmatrix} = (-5-9)\mathbf{i} - (15-9)\mathbf{j} + (3+1)\mathbf{k} = [-14, -6, 4].$$

The signed volume of the parallelepiped with vertices in O, P, Q, R is (cf. (1.11)):

$$[\overrightarrow{OP}, \overrightarrow{OQ}, \overrightarrow{OR}] = \begin{vmatrix} 1 & 2 & 4 \\ -2 & 3 & -5 \\ 0 & 1 & -1 \end{vmatrix} = 1(-3+5) + 2(-2-4) = 2 - 12 = -10.$$

Remark that we used expansion by *minors* on the first *column* to calculate the determinant.

1.2.6. *Formal properties of products.* The following properties are stated without proof:

- $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$;
- $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = (\mathbf{x} \cdot \mathbf{y}) + (\mathbf{x} \cdot \mathbf{z})$;
- $a\mathbf{x} \cdot \mathbf{y} = a(\mathbf{x} \cdot \mathbf{y})$;
- $[\mathbf{y}, \mathbf{x}] = -[\mathbf{x}, \mathbf{y}]$;
- $\mathbf{x} \cdot \hat{\mathbf{y}} = [\mathbf{y}, \mathbf{x}] = -[\mathbf{x}, \mathbf{y}]$;
- $\mathbf{y} \times \mathbf{x} = -(\mathbf{x} \times \mathbf{y})$;
- $\mathbf{x} \times (\mathbf{y} + \mathbf{z}) = (\mathbf{x} \times \mathbf{y}) + (\mathbf{x} \times \mathbf{z})$;
- $a\mathbf{x} \times \mathbf{y} = a(\mathbf{x} \times \mathbf{y})$;
- In general, $\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) \neq (\mathbf{x} \times \mathbf{y}) \times \mathbf{z}$;
- $[\mathbf{x} + \mathbf{x}', \mathbf{y}, \mathbf{z}] = [\mathbf{x}, \mathbf{y}, \mathbf{z}] + [\mathbf{x}', \mathbf{y}, \mathbf{z}]$;

- $[a\mathbf{x}, \mathbf{y}, \mathbf{z}] = a[\mathbf{x}, \mathbf{y}, \mathbf{z}]$.
- $[\mathbf{x}, \mathbf{y}, \mathbf{z}] = -[\mathbf{y}, \mathbf{x}, \mathbf{z}] = -[\mathbf{x}, \mathbf{z}, \mathbf{y}]$.

2. Description of Lines and Planes

Next to points, the most elementary geometric objects are the *lines* in plane $\mathbf{E}^2$ and *lines* and *planes* in space $\mathbf{E}^3$. In this section, we want to give effective descriptions of lines and planes and describe how to use these concepts to answer geometrical questions. We fix a point O as origin and a coordinate system $\{\mathbf{i}, \mathbf{j}\}$ for $\mathbf{R}^2$, resp. $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ for $\mathbf{R}^3$. In most cases, definitions and properties resemble each other in the plane, resp. the space case. We write $\mathbf{R}^i$, resp. $\mathbf{E}^i$ to cover both cases. So either $i = 2$ or $i = 3$.

2.1. Parametrizations.

2.1.1. *Lines.* A line l in the plane $\mathbf{E}^2$ or in space $\mathbf{E}^3$ is determined by the following property: There is a non-zero vector $\mathbf{x} \in \mathbf{R}^i$ such that

$$\{\overrightarrow{PQ} | P, Q \in l\} = sp(\mathbf{x}).$$

Any vector $\mathbf{x}$ with this property is called a *parallel vector* for l. In particular, every vector joining two different points on l is a parallel vector; in fact, it is a certain multiple $t\mathbf{x}$, $t \in \mathbf{R}$. On the other hand, given any point $P \in l$, then every point $Q \in \mathbf{E}^i$ with $\overrightarrow{PQ} = t\mathbf{x}$ is on the line l. This leads to the following definition of a parametrization of a line:

DEFINITION 1.9. Given a point $P \in \mathbf{E}^i$ and a non-zero vector $\mathbf{x} \in \mathbf{R}^i$. The *line through* P with *parallel vector* $\mathbf{x}$ consists of *all* points $Q \in \mathbf{E}^i$ with $\overrightarrow{PQ} \in sp(\mathbf{x})$, or equivalently, such that

$$\overrightarrow{OQ} = \overrightarrow{OP} + t\mathbf{x}, \quad t \in \mathbf{R}. \tag{1.14}$$

One way to imagine the parametrization (1.14) is to think of a particle that moves with constant speed on the line l. Then t is a time parameter, and the particle starts at P at time $t = 0$.

Given two distinct points $P, P' \in \mathbf{E}^i$. Then there is precisely one line l containing both P and P'. Its parametrization is given by

$$l = \{Q \in \mathbf{E}^i | \overrightarrow{OQ} = \overrightarrow{OP} + t\overrightarrow{PP'},\ t \in \mathbf{R}\}.$$

The *line segment* of all points on l between P and P' consists of all points

$$l = \{Q \in \mathbf{E}^i | \overrightarrow{OQ} = \overrightarrow{OP} + t\overrightarrow{PP'},\ t \in [0, 1]\}.$$

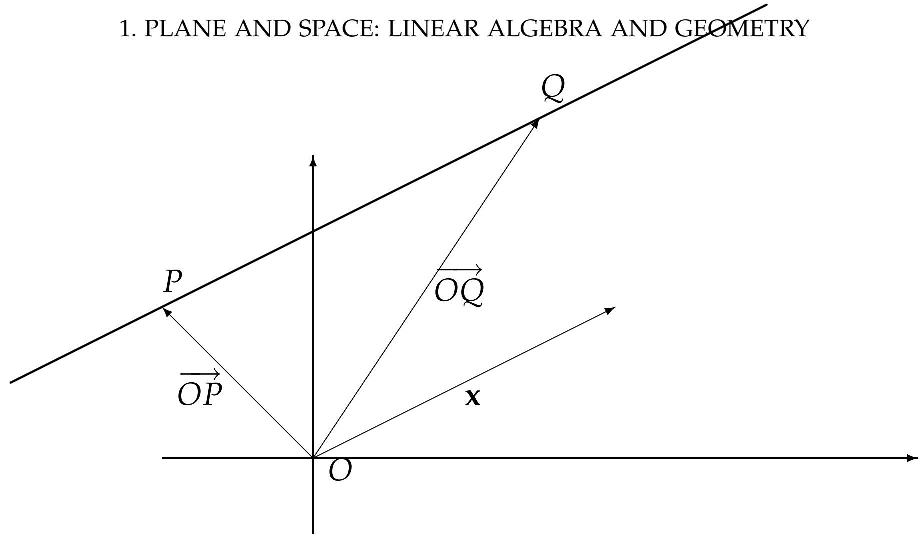


FIGURE 5. The line l through P with parallel vector $\mathbf{x}$

EXAMPLE 1.10.
A parametrization for the line l through $P[1,2,-2]$ and $P'[3,1,3]$ is found as follows:

$$\mathbf{x} = \overrightarrow{PP'} = [3-1, 1-2, 3-(-2)] = [2,-1,5];$$

$$l = \{Q \in \mathbf{E}^3 | \overrightarrow{OQ} = [1,2,-2] + t[2,-1,5],\ t \in \mathbf{R}\}.$$

The point $Q_1 = [-3,4,-12]$ is on the line l and corresponds to $t = -2$, whereas $Q_2 = [2,4,6]$ is not on l, since $\overrightarrow{PQ_2} = [1,2,8] \notin sp([2,-1,5])$.

REMARK 1.11.
A parametrization of a line l is *not unique*. In fact, the same line l has *infinitely many* linear parametrizations. Here is how to get them from the one given in (1.14): You may replace $\mathbf{x}$ by any non-zero vector $\mathbf{x}'$ parallel to $\mathbf{x}$, and, at the same time, you may replace P by any other point $P' \in l$.

2.1.2. *Planes.* A plane α in space $\mathbf{E}^3$ is determined by the following property: There are two *linearly independent* vectors $\mathbf{x}, \mathbf{y} \in \mathbf{R}^3$ such that

$$\{\overrightarrow{PQ} | P, Q \in \alpha\} = sp(\mathbf{x}, \mathbf{y}).$$

In particular, every *parallel vector* joining two points on α is some linear combination $s\mathbf{x} + t\mathbf{y}, \quad s, t \in \mathbf{R}$. Conversely, given any point $P \in \alpha$, then every point $Q \in \mathbf{E}^3$ with $\overrightarrow{PQ} = s\mathbf{x} + t\mathbf{y}$ is in the plane α. This leads to the following definition of a parametrization of a plane:

DEFINITION 1.12. Given a point $P \in \mathbf{E}^3$ and two linearly independent vectors $\mathbf{x}, \mathbf{y} \in \mathbf{R}^3$. The *plane through P* with *parallel plane* $sp(\mathbf{x}, \mathbf{y})$ consists of all points $Q \in \mathbf{E}^3$ with $\overrightarrow{PQ} \in sp(\mathbf{x}, \mathbf{y})$, or equivalently, such that

$$\overrightarrow{OQ} = \overrightarrow{OP} + s\mathbf{x} + t\mathbf{y}, \quad s, t \in \mathbf{R}. \tag{1.15}$$

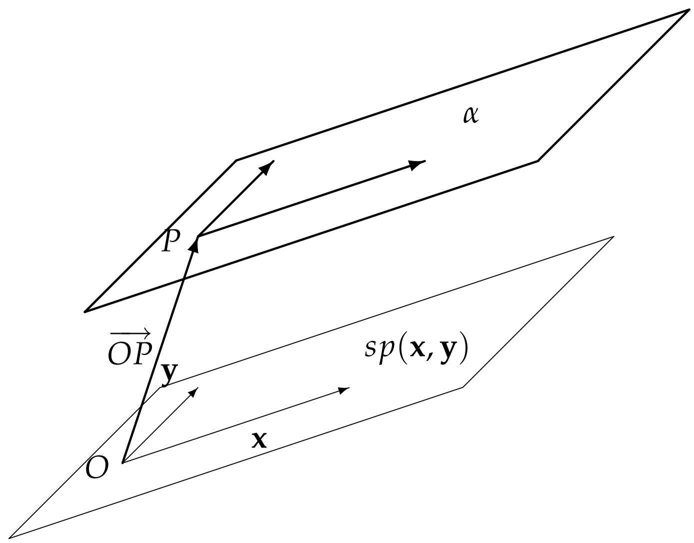


FIGURE 6. The plane α through P with parallel plane $sp(\mathbf{x}, \mathbf{y})$

Given three distinct points $P_1, P_2, P_3 \in \mathbf{E}^3$. Then the set

$$\{Q \in \mathbf{E}^3 | \overrightarrow{OQ} = \overrightarrow{OP_1} + s\overrightarrow{P_1P_2} + t\overrightarrow{P_1P_3}, \quad s, t \in \mathbf{R}\}$$

- is a line, if $\overrightarrow{P_1P_2}$ and $\overrightarrow{P_2P_3}$ are parallel;
- is *the* plane containing P_1, P_2, P_3 otherwise.

To decide whether a given point $Q \in \mathbf{E}^3$ is contained in the plane α with parametrization $\overrightarrow{OQ} = \overrightarrow{OP} + s\mathbf{x} + t\mathbf{y}$ from (1.15) amounts to considering (1.15) as a *vector equation*, or equivalently, as a *system* of three *linear equations* in the two parameters s and t.

EXAMPLE 1.13.
Let $P_1 = [1, 1, 1]$, $P_2 = [2, 2, 1]$, $P_3 = [3, 2, 2]$, and $Q = [1, -2, 4]$. Then $\overrightarrow{P_1P_2} = [1, 1, 0]$, $\overrightarrow{P_1P_3} = [2, 1, 1]$, and

$$\alpha = \{[x, y, z] \in \mathbf{E}^3 | \ [x, y, z] = [1, 1, 1] + s[1, 1, 0] + t[2, 1, 1], \quad s, t \in \mathbf{R}\}$$

is a parametrization for the plane through P_1, P_2, P_3 – since $[1,1,0]$ and $[2,1,1]$ are linearly independent. To decide whether $Q \in \alpha$, we must find out whether the equation

$$\overrightarrow{P_1Q} = \overrightarrow{OQ} - \overrightarrow{OP_1} = s\overrightarrow{P_1P_2} + t\overrightarrow{P_1P_3}$$

has a solution. In our case, one has to solve the equation

$$[0,-3,3] = s[1,1,0] + t[2,1,1],$$

or equivalently the system

$$\begin{aligned} s+2t &= 0 \\ s+t &= -3 \\ t &= 3. \end{aligned}$$

Using Gauss-Jordan reduction, we get:

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 1 & 1 & -3 \\ 0 & 1 & 3 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & -1 & -3 \\ 0 & 1 & 3 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -6 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right].$$

We end with a solution $s = -6$, $t = 3$, and hence $Q \in \alpha$. A similar calculation replacing only Q with $Q'[1,-2,3]$ ends with a matrix whose last row is $[0\ 0| -1]$. Hence, that new system has no solution, and $Q' \notin \alpha$.

REMARK 1.14.
A parametrization of a given plane is *not* unique: The plane α through P with parallel plane $sp(\mathbf{x}, \mathbf{y})$ is equal to the plane α' through P' with parallel plane $sp(\mathbf{x}', \mathbf{y}')$ if and only if $sp(\mathbf{x}, \mathbf{y}) = sp(\mathbf{x}', \mathbf{y}')$ and if P' is contained in α.

2.2. Equations. An alternative way to describe lines in the plane $\mathbf{E}^2$ or planes in space $\mathbf{E}^3$ is by a *linear equation* in the two variables x and y, resp. the three variables x, y and z. Deriving these equations, we make use of *orthogonality* using the *dot product*, cf. (1.4).

2.2.1. *Lines in the plane.* Given a line l in Euclidean plane $\mathbf{E}^2$ with parallel vector $\mathbf{x}$. A vector $\mathbf{n} \neq \mathbf{0}$ is called a *normal vector* to l, if it is orthogonal to the parallel vector $\mathbf{x}$, i.e., if $\mathbf{n} \cdot \mathbf{x} = 0$. In particular, $\mathbf{n}$ has to be some non-zero multiple of the hat vector $\hat{\mathbf{x}}$.

Now let us derive an equation describing the line l through the point $P \in \mathbf{E}^2$ with normal vector $\mathbf{n}$. Let $Q \in \mathbf{E}^2$ denote an arbitrary point. To check whether $Q \in l$, we proceed as follows:

$$\begin{aligned} Q \in l &\Leftrightarrow \mathbf{x}, \overrightarrow{PQ} \text{ are parallel} \\ &\Leftrightarrow \mathbf{n} \cdot \overrightarrow{PQ} = 0 \\ &\Leftrightarrow \mathbf{n} \cdot (\overrightarrow{OQ} - \overrightarrow{OP}) = 0 \\ &\Leftrightarrow \mathbf{n} \cdot \overrightarrow{OQ} = \mathbf{n} \cdot \overrightarrow{OP}. \end{aligned} \tag{1.16}$$

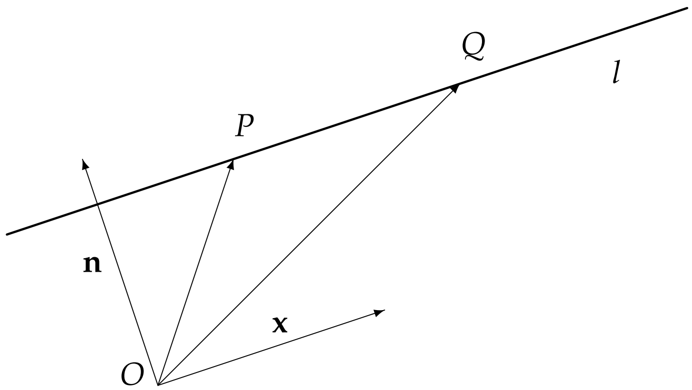


FIGURE 7. Derivation of an equation describing a line l

In coordinates, equation (1.16) reads as follows:
Let $\overrightarrow{OQ} = [x, y], \overrightarrow{OP} = [x_0, y_0]$ and $\mathbf{n} = [a, b]$. Then, (1.16) is equivalent to

$$a(x - x_0) + b(y - y_0) = 0 \Leftrightarrow ax + by = ax_0 + by_0 = d \tag{1.17}$$

with d the fixed number $d = \mathbf{n} \cdot \overrightarrow{OP}$. Hence, $Q[x, y] \in l$ is equivalent to (1.17).

EXAMPLE 1.15.
Given $P_1[2, 5]$, $P_2[-1, 4] \in \mathbf{E}^2$. An equation for the line through P_1 and P_2 is determined as follows:

$$\begin{aligned} \mathbf{x} = \overrightarrow{P_1P_2} &= [-3, -1]; \\ \mathbf{n} = \hat{\mathbf{x}} &= [1, -3]; \\ \mathbf{n} \cdot \overrightarrow{OP_1} &= [1, -3] \cdot [2, 5] = 2 - 15 = -13 \end{aligned}$$

The line l can thus be described by the equation $x - 3y = -13$. You may check that the coordinates of P_1 and P_2 actually satisfy that equation.

2.2.2. *Planes in space.* For a plane $\alpha \subset \mathbf{E}^3$ with parallel plane $sp(\mathbf{x}, \mathbf{y})$, a vector $\mathbf{n}$ is called a *normal vector* to α if $\mathbf{n}$ is orthogonal to every vector in the parallel plane. In particular, $\mathbf{n}$ has to be some non-zero multiple of the cross product vector $\mathbf{x} \times \mathbf{y}$.

For a plane α through $P \in \mathbf{E}^3$ with normal vector $\mathbf{n}$, we obtain an equation describing α using that normal vector as follows:

$$\begin{aligned} Q \in \alpha &\Leftrightarrow \mathbf{n} \text{ and } \overrightarrow{PQ} \text{ are perpendicular} \\ &\Leftrightarrow \mathbf{n} \cdot \overrightarrow{PQ} = 0. \end{aligned} \tag{1.18}$$

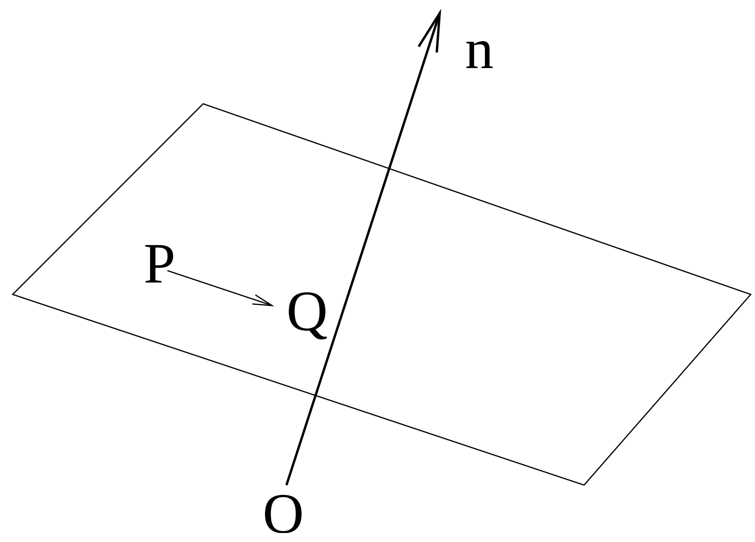


FIGURE 8. Plane and normal vector

Using $\mathbf{n} = \mathbf{x} \times \mathbf{y}$ as normal vector, we obtain the following equivalent conditions:

$$Q \in \alpha \quad \Leftrightarrow \quad (\mathbf{x} \times \mathbf{y}) \cdot \overrightarrow{PQ} = 0$$

$$\Leftrightarrow \quad [\mathbf{x}, \mathbf{y}, \overrightarrow{PQ}] = 0 \tag{1.19}$$

$$\Leftrightarrow \quad (\mathbf{x} \times \mathbf{y}) \cdot \overrightarrow{OQ} = (\mathbf{x} \times \mathbf{y}) \cdot \overrightarrow{OP} \tag{1.20}$$

$$\Leftrightarrow \quad [\mathbf{x}, \mathbf{y}, \overrightarrow{OQ}] = [\mathbf{x}, \mathbf{y}, \overrightarrow{OP}]. \tag{1.21}$$

To see that all these correspond to a linear equation, let $\overrightarrow{OQ} = [x, y, z], \overrightarrow{OP} = [x_0, y_0, z_0]$ and $\mathbf{n} = \mathbf{x} \times \mathbf{y} = [a, b, c]$. Remark that x, y, z are variables, while a, b, c, x_0, y_0, z_0 are fixed real numbers. Then, (1.18) is equivalent to

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \Leftrightarrow ax + by + cz = ax_0 + by_0 + cz_0 = d, \tag{1.22}$$

with d the fixed number $d = \mathbf{n} \cdot \overrightarrow{OP} = (\mathbf{x} \times \mathbf{y}) \cdot \overrightarrow{OP} = [\mathbf{x}, \mathbf{y}, \overrightarrow{OP}]$. Hence, $Q[x, y, z] \in \alpha$ if its coordinates satisfy equation (1.22).

Remark the following geometric interpretation for (1.19): The parallelepiped spanned by $\mathbf{x}, \mathbf{y}, \overrightarrow{PQ}$ has 0 volume (cf. 1.2.5) if and only if $\overrightarrow{PQ}$ is contained in the plane spanned by $\mathbf{x}$ and $\mathbf{y}$.

EXAMPLE 1.16.
Let $P_1, P_2, P_3 \in \mathbf{E}^3$ denote three points in space that are *not* contained in a common line. We want to derive an equation, which determines the plane α containing these three points using the space product as in (1.19):

Let $\overrightarrow{OQ} = [x, y, z]$, $\overrightarrow{OP_i} = [x_i, y_i, z_i]$, $1 \leq i \leq 3$. Then,

$$\mathbf{x} = \overrightarrow{P_1P_2} = [x_2 - x_1, y_2 - y_1, z_2 - z_1] \text{ and } \mathbf{y} = \overrightarrow{P_1P_3} = [x_3 - x_1, y_3 - y_1, z_3 - z_1]$$

span a parallel plane to α, and $\overrightarrow{P_1Q} = [x - x_1, y - y_1, z - z_1]$. Hence

$$\begin{aligned} Q \in \alpha \quad &\Leftrightarrow \quad [\mathbf{x}, \mathbf{y}, \overrightarrow{P_1Q}] = 0 \\ &\Leftrightarrow \quad \begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \\ x - x_1 & y - y_1 & z - z_1 \end{vmatrix} = 0. \end{aligned}$$

EXAMPLE 1.17.
Let us now look at the particular case $P_1[1, 2, 1]$, $P_2[-1, 2, 3]$, $P_3[2, 1, 4] \in \mathbf{E}^3$. An equation determining the plane α containing these three points is determined as follows:

$$0 = \begin{vmatrix} -2 & 0 & 2 \\ 1 & -1 & 3 \\ x-1 & y-2 & z-1 \end{vmatrix} = (x-1)2 - (y-2)(-8) + (z-1)2 = 2x + 8y + 2z - 20.$$

Substituting the coordinates of any of the three points P_i into x, y and z allows you to check that this is a correct equation for α.

REMARK 1.18.
You may ask yourself whether *any* equation $ax + by = d$ characterizes a line in the plane or whether *any* equation $ax + by + cz = d$ characterizes a plane in space:

(1) It is not difficult to see, that an equation $ax + by = d$ (for $[a, b] \neq \mathbf{0}$) is solved by the coordinates $[x, y]$ of points Q on a line l in the plane $\mathbf{E}^2$: In fact, the vector $\mathbf{n} = [a, b]$ may serve as a normal vector to l, and the line can always be parametrized in the following way:

$$l = \{Q \in \mathbf{E}^2 | \overrightarrow{OQ} = [e, f] + t[-b, a], \quad t \in \mathbf{R}\}$$

with $[e, f]$ any vector solving the equation $ae + bf = d$. For $a \neq 0$, one may use $[e, f] = [\frac{d}{a}, 0]$, for $b \neq 0$, a simple solution has coordinates $[0, \frac{d}{b}]$. A solution that always works is $[e, f] = [\frac{ad}{a^2+b^2}, \frac{bd}{a^2+b^2}]$. Note that the parallel vector $[-b, a]$ is the hat vector to the vector $\mathbf{n} = [a, b]$.

(2) Likewise, the equation $ax + by + cz = d$ (for $[a, b, c] \neq \mathbf{0}$) is solved by the coordinates $[x, y, z]$ of points Q on a plane α in space $\mathbf{E}^3$ with parametrization

$$\alpha = \{Q \in \mathbf{E}^2 | \overrightarrow{OQ} = [e, f, g] + s\mathbf{x} + t\mathbf{y}, \quad s, t \in \mathbf{R}\}$$

where $[e, f, g]$ can be one of the (solution) vectors $[\frac{d}{a}, 0, 0], [0. \frac{d}{b}, 0]$ or $[0, 0, \frac{d}{c}]$ – choose one with denominator different from 0! – or

$[e, f, g] = [\frac{ad}{a^2+b^2+c^2}, \frac{bd}{a^2+b^2+c^2}, \frac{cd}{a^2+b^2+c^2}]$. Moreover, you may choose any two non-zero vectors $\mathbf{x}, \mathbf{y}$ among $[-b, a, 0]$, $[-c, 0, a]$, $[0, -c, b]$ to span a parallel plane. Why? An easy dot-product calculation shows that they are normal to $\mathbf{n} = [a, b, c]$. Hence, they may serve as (non-unique) replacements in space for the plane hat-vector!

(3) The linear equation representing a line in the plane, resp. a plane in space is unique up to a non-zero factor: The equations $ax + by = d$ and $tax + tby = td$, $t \in \mathbf{R} \setminus \{0\}$, have the same solution. Thus, the equation from Ex. 1.17 can be replaced by $x + 4y + z = 10$.

(4) A line in space may be given as the set of solutions of *two* linear equations in three unknowns x, y, z, corresponding to the intersection of the two planes that each of the equations represents. More about this topic follows subsequently!

2.3. Several Lines or Planes. In this section, we discuss the *relative* position of several lines in the plane and in space, resp. of several planes in space. Do they intersect, are they prallel etc.?

2.3.1. *Two lines in the plane.*

PROPOSITION 1.19. *The intersection of two lines l_1, l_2 in the Euclidean plane $\mathbf{E}^2$ is either*

(1) *a single point S;*
(2) *empty, or*
(3) $l_1 = l_2$.

In case 2. *and* 3.*, the lines are called* parallel *to each other* ($l_1 \parallel l_2$).

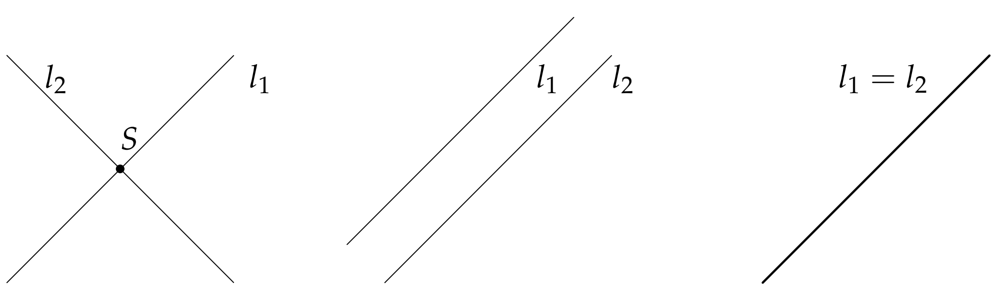


FIGURE 9. Intersection of two lines

Suppose, l_1 and l_2 are given by parametrizations, resp. by equations. How can we distinguish the three cases, and, in case 1., how can we calculate the point S of intersection? Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, and suppose

$$\begin{aligned} l_1 &= \{Q \in \mathbf{E}^2 | \overrightarrow{OQ} = \overrightarrow{OP} + s\mathbf{x}, \quad s \in \mathbf{R}\}; \\ l_2 &= \{R \in \mathbf{E}^2 | \overrightarrow{OR} = \overrightarrow{OP'} + t\mathbf{y}, \quad t \in \mathbf{R}\}. \end{aligned}$$

Then, $l_1 \cap l_2$ consists of those $S \in \mathbf{E}^2$, that satisfy both equations:

$$\overrightarrow{OS} = \overrightarrow{OP} + s\mathbf{x} = \overrightarrow{OP'} + t\mathbf{y}. \tag{1.23}$$

The solutions s, t of the vector equation (1.23) – if existing – have to satisfy

$$s\mathbf{x} - t\mathbf{y} = \overrightarrow{OP'} - \overrightarrow{OP} = \overrightarrow{PP'}.$$

Coordinatewise, this last vector equation corresponds to two linear equations in the two variables s and t, or equivalently, to the matrix equation

$$\mathbf{A} \begin{bmatrix} s \\ -t \end{bmatrix} = \overrightarrow{PP'} \text{ with } \mathbf{A} = [\mathbf{x}\,\mathbf{y}] = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}. \tag{1.24}$$

Then, the three cases in Prop. 1.19 correspond to:

(1) $\mathbf{x}$ and $\mathbf{y}$ are linearly independent, i.e., the rank of the matrix $\mathbf{A}$ is two (rank $\mathbf{A}$=2), i.e., $\mathbf{A}$ can be reduced to the identity matrix $\mathbf{I}$ by the *Gauss-Jordan algorithm* from linear algebra. This algorithm, applied to the *extended* matrix $[\mathbf{A}|\overrightarrow{PP'}]$ yields $\left[\mathbf{I}|\begin{bmatrix} s_Q \\ -t_Q \end{bmatrix}\right]$. Now, the coordinates of the point Q of intersection are found by introducing s_Q at the place of s in (1.23) (or t_Q at the place of t. In fact, it is only necessary to calculate either s_Q or t_Q).

(2) $\mathbf{x}$ and $\mathbf{y}$ are parallel, but $\mathbf{x}$ and $\overrightarrow{PP'}$ are linearly independent. This corresponds to: rank $\mathbf{A}$=1, and the Gauss-Jordan algorithm reduces $[\mathbf{A}|\overrightarrow{PP'}]$ to a matrix whose second row is $[00|t]$ with $t \neq 0$. Thus, the matrix equation (1.24) does not have any solution, i.e., $l_1 \cap l_2 = \emptyset$.

(3) $\mathbf{x}$, $\mathbf{y}$ and $\overrightarrow{PP'}$ are parallel. This corresponds to: rank $\mathbf{A}$=1, and the Gauss-Jordan algorithm reduces $[\mathbf{A}|\overrightarrow{PP'}]$ to a matrix whose second row is $[00|0]$. Hence the matrix equation (1.24) is equivalent to the solution of the equation determining just l_1, i.e., $l_1 = l_1 \cap l_2 = l_2$.

EXAMPLE 1.20.
Let $P[5,-1]$, $P'[-7,3]$, $\mathbf{x} = [1,-2]$, $\mathbf{y} = [3,-7]$. Then, $\overrightarrow{PP'} = [-12,4]$, and the Gauss-Jordan algorithm applied to $[\mathbf{A}|\overrightarrow{PP'}]$ yields:

$$[\mathbf{A}|\overrightarrow{PP'}] = \left[\begin{array}{rr|r} 1 & 3 & -12 \\ -2 & -7 & 4 \end{array}\right] \simeq \left[\begin{array}{rr|r} 1 & 3 & -12 \\ 0 & -1 & -20 \end{array}\right] \simeq \left[\begin{array}{rr|r} 1 & 0 & -72 \\ 0 & 1 & 20 \end{array}\right].$$

With $s_Q = -72$, we obtain the point Q of intersection by
$\overrightarrow{OS} = \overrightarrow{OP} + s_Q\mathbf{x} = [5,-1] + [-72,144] = [-67,143]$.

If the two lines l_1, l_2 instead are given by the linear equations

$$\begin{array}{rcl} a_1x + b_1y & = & d_1 \\ a_2x + b_2y & = & d_2, \end{array}$$

the solutions $[x,y]$ of this system correspond to points of intersection. In this case one has to perform a Gauss-Jordan reduction to the extended matrix $\left[\begin{array}{cc|c} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \end{array}\right]$.

2.3.2. *Two lines in space.* In Euclidean space $\mathbf{E}^3$, we observe a *new phenomenon*:

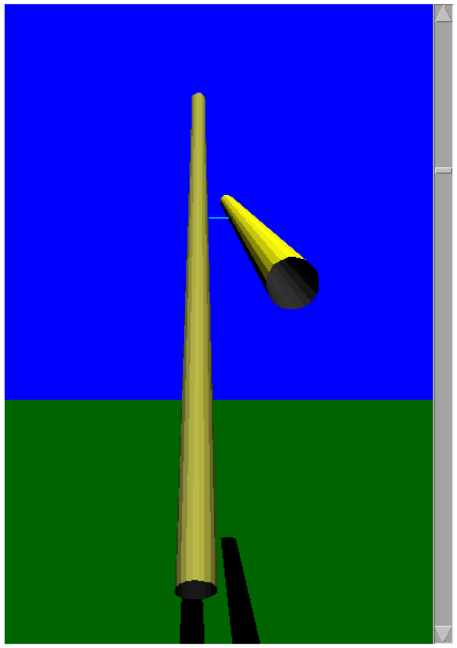
DEFINITION 1.21.
(1) Two lines $l_1, l_2 \subset \mathbf{E}^3$ are called *skew*[a], if there is *no* plane $\alpha \subset \mathbf{E}^3$ containing both of them. (Note, as a consequence: l_1 and l_2 do *not* intersect!)
(2) Two lines $l_1, l_2 \subset \mathbf{E}^3$ are called *parallel*, ($l_1 \parallel l_2$) if there is a plane $\alpha \subset \mathbf{E}^3$ containing both of them and such that $l_1 \cap l_2$ is empty or $l_1 = l_2$.

[a] illustration: opposite page

EXAMPLE 1.22.
Consider l_1 to be the X-axis, and l_2 to be a parallel to the Z-axis and not intersecting the X-axis. Verify that l_1 and l_2 are skew!

Hence, in space one has to consider four cases: Two lines may be contained in a common plane or not; if not, they are skew. If they are contained in a common plane, one distinguishes again the three cases from Prop. 1.19 and proceeds in the same way as in (1.23) ff.: With $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, the vector equation $s\mathbf{x} - t\mathbf{y} = \overrightarrow{PP'}$

Here are two skew lines:



Rotate the two line segments with the mouse and zoom in or out using the scrollbar.

FIGURE 10. Skew lines

corresponds to a matrix equation

$$\mathbf{A}\begin{bmatrix} s \\ -t \end{bmatrix} = \overrightarrow{PP'} \quad \text{with} \quad \mathbf{A} = [\mathbf{x}\,\mathbf{y}] = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix},$$

i.e., to a system of *three* linear equations in the two parameter values s and t. The new case – corresponding to skew lines l_1 and l_2 – occurs, whenever $\mathbf{x}, \mathbf{y}$ and $\overrightarrow{PP'}$ are linearly *in*dependent. In that case, the matrix $[\mathbf{A}|\overrightarrow{PP'}]$ can be reduced to a matrix $\left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & b \\ \hline 0 & 0 & c \end{array}\right]$ with $c \neq 0$. The last equation associated to the latter extended matrix is: $0s + 0t = c$, and it has obviolusly no solution: The two lines l_1 and l_2 do not intersect!

2.3.3. *Two planes in space.*

PROPOSITION 1.23. *The intersection of two planes α_1, α_2 in Euclidean space $\mathbf{E}^3$ is either*
(1) *a line l;*
(2) *empty, or*
(3) $\alpha_1 = \alpha_2$.
In case 2. *and* 3.*, the planes are called* parallel *to each other* ($\alpha_1 \parallel \alpha_2$).

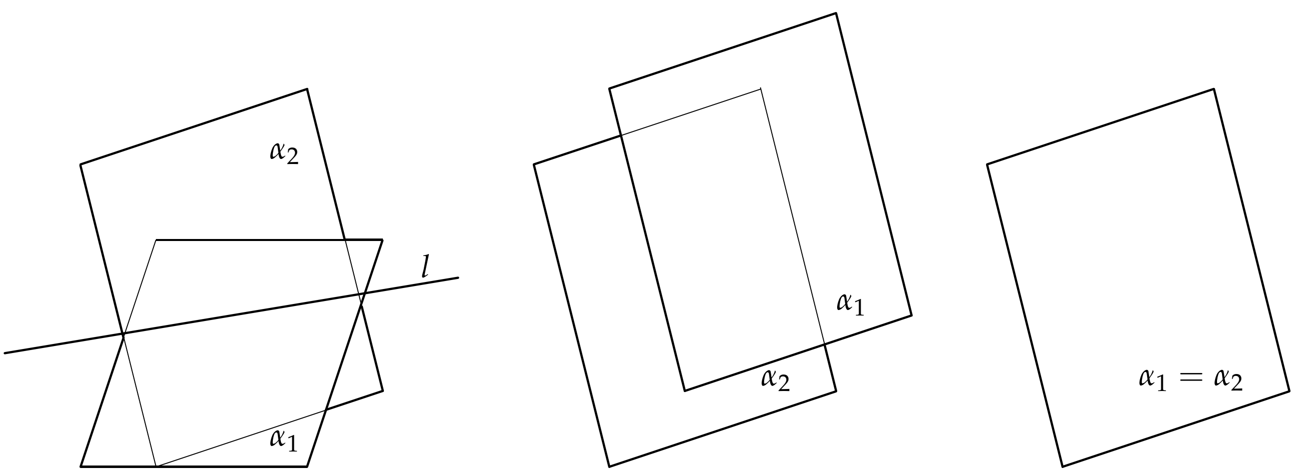


FIGURE 11. Intersection of two planes

It is easiest to find the intersection of two planes $\alpha_1, \alpha_2 \subset \mathbf{E}^3$ when both are given by linear equations:

$$\begin{aligned} a_1 x + b_1 y + c_1 z &= d_1 \\ a_2 x + b_2 y + c_2 z &= d_2. \end{aligned} \tag{1.25}$$

Solve this system of equations by applying the Gauss-Jordan algorithm to the extended matrix $[\mathbf{A}|\mathbf{d}]$ with $\mathbf{A} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix}$ and $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$. If rank $\mathbf{A} = 2$, one of the variables is *free* and can be used as a parameter representing the line of intersection. If rank $\mathbf{A} = 1$, the system has either *no* solution (parallel planes) or both equations determine the same plane.

EXAMPLE 1.24.
Let α_1, α_2 be given by

$$\begin{aligned} x - y + 2z &= 1 \\ 2x + y - 2z &= 2. \end{aligned} \tag{1.26}$$

The Gauss-Jordan algorithm transforms $\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 2 & 1 & -2 & 2 \end{array}\right]$ into $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & -2 & 0 \end{array}\right]$.
Using $z = t$ as the free variable, this yields the solution $[x, y, z] = [1, 2t, t] = [1, 0, 0] + t[0, 2, 1],\ t \in \mathbf{R}$, which in fact is a parametrization for the line $l = \alpha_1 \cap \alpha_2$.

2.3.4. *A line and a plane in space.* Finally, one may consider the intersection of a line l and a plane α in space. If both are given by a parametrization, one may determine their intersection as in 2.3.1; this time, one obtains a system of three linear equations in *three* variables - the parameters for both l and α. If the coefficient matrix $\mathbf{A}$ corresponding to this system has rank 3, there is a *unique* solution, corresponding to a point Q of intersection. If rank $\mathbf{A}$=2, then l is parallel to α, ($l \parallel \alpha$). As a special case, one may obtain: $l \subset \alpha$.

If the plane α is given by a linear equation $ax + by + cz = d$, we insert the components of a parametrization $\overrightarrow{OQ} = \overrightarrow{OP} + t\mathbf{x}$ of the line l into that equation and obtain a *single* linear equation in the variable t. A solution t of the latter inserted into the parametrization for l yields the vector $\overrightarrow{OQ}$ with Q the point of intersection.

EXAMPLE 1.25.
Let the plane α be given by the equation $x - 2y + z = 4$, and the line l be given by the parametrization $l = \{Q \in \mathbf{E}^3 | \overrightarrow{OQ} = [x, y, z] = [1, 2, -2] + t[2, -1, 5],\ t \in \mathbf{R}\}$. Then the parameter t corresponding to the unique point $Q \in l \cap \alpha$ has to satisfy:
$4 = x - 2y + z = 1 + 2t - 2(2 - t) + (-2 + 5t)$, hence $-5 + 9t = 4$ and $t = 1$.
Inserting $t = 1$ into the parametrization, one obtains
$\overrightarrow{OQ} = [1, 2, -2] + [2, -1, 5] = [3, 1, 3]$, and $Q[3, 1, 3]$ is the unique point of intersection.

3. Orthogonal Projections, Distances and Angles

3.1. Orthogonal Projections. An orthogonal projection gives us an easy way of decomposing the plane or space wrt. an embedded line or plane; both dot and cross products are used in calculations. We use projections in particular to define and calculate *distances* and *angles*. How would you for example determine the distance between two (skew) lines?

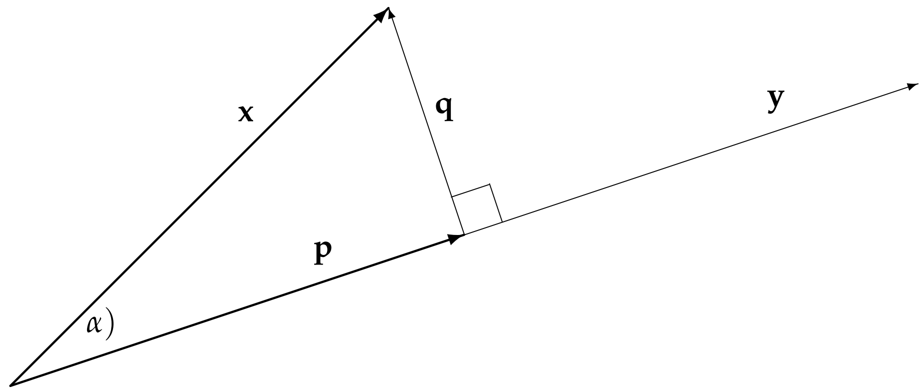


FIGURE 12. Orthogonal projection of $\mathbf{x}$ on $\mathbf{y}$

3.1.1. *Orthogonal projections of vectors.* In many applications of vector geometry, one would like to find out which part of a given vector $\mathbf{x}$ can be considered to act in the direction given by another vector $\mathbf{y} \neq \mathbf{0}$. Think of the graviatational force vector) acting on an object placed on an incline. What is the component of this force in direction of the incline?

It is our aim, to decompose $\mathbf{x}$ into a vector $\mathbf{p}$ *parallel* to $\mathbf{y}$ and a vector $\mathbf{q}$ *orthogonal* to $\mathbf{y}$, i.e.,

$$\mathbf{x} = \mathbf{p} + \mathbf{q} \text{ with } \mathbf{p} = a\mathbf{y} \text{ and } \mathbf{q} \cdot \mathbf{y} = 0. \tag{1.27}$$

The vector $\mathbf{p}$ is then called the *orthogonal projection* of $\mathbf{x}$ on $\mathbf{y}$.

PROPOSITION 1.26. *The orthogonal projection of the vector* $\mathbf{x}$ *on the line* $sp(\mathbf{y})$ *in* $\mathbf{R}^i$, $i = 2$ *or* $i = 3$, *is given by*

$$\blacksquare \quad \mathbf{p} = \left(\frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{y} \cdot \mathbf{y}}\right) \mathbf{y}. \quad \blacksquare \tag{1.28}$$

The vector $\mathbf{p}$ *is parallel to* $\mathbf{y}$ *and has length* $|\mathbf{p}| = \frac{|\mathbf{x} \cdot \mathbf{y}|}{|\mathbf{y}|} = |\mathbf{x}| \cos \alpha$, *where* α *denotes the angle between x and y.*

PROOF:
Performing the dot product with $\mathbf{y}$ to both sides of (1.27) yields: $\mathbf{x} \cdot \mathbf{y} = \mathbf{p} \cdot \mathbf{y} = a\mathbf{y} \cdot \mathbf{y}$, and hence $a = \frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{y} \cdot \mathbf{y}}$. The length $|\mathbf{p}|$ can be derived from the definition of the cosine function in the rightangled triangle in Fig. 12 or from the following calculation:

$$|\mathbf{p}|^2 = \mathbf{p} \cdot \mathbf{p} = \mathbf{p} \cdot (\mathbf{x} - \mathbf{q}) = \mathbf{p} \cdot \mathbf{x} = |\mathbf{p}||\mathbf{x}| \cos \alpha.$$

□

COROLLARY 1.27. *The orthogonal vector* $\mathbf{q}$ *in decomposition* (1.27) *above is given as*

$$\text{(1.29)} \quad \textit{(plane)}\ \mathbf{q} = \left(\frac{[\mathbf{y},\mathbf{x}]}{\mathbf{y}\cdot\mathbf{y}}\right)\hat{\mathbf{y}} \textit{ with length } |\mathbf{q}| = |\mathbf{x}|\sin\alpha = \frac{|[\mathbf{y},\mathbf{x}]|}{|\mathbf{y}|};$$

$$\text{(1.30)} \quad \textit{(space)}\ \mathbf{q} = \mathbf{y}\times\frac{(\mathbf{x}\times\mathbf{y})}{\mathbf{y}\cdot\mathbf{y}} \textit{ with length } |\mathbf{q}| = |\mathbf{x}|\sin\alpha = \frac{|\mathbf{x}\times\mathbf{y}|}{|\mathbf{y}|}.$$

REMARK 1.28.
If one wants to determine *both* $\mathbf{p}$ and $\mathbf{q}$, it is often easier to use:
$\mathbf{q} = \mathbf{x} - \mathbf{p}$ and $|\mathbf{q}|^2 = |\mathbf{x}|^2 - |\mathbf{p}|^2$.

PROOF. Let us first find a vector $\mathbf{z}$ that is perpendicular to $\mathbf{y}$ and contained in the plane spanned by $\mathbf{x}$ and $\mathbf{y}$. In the plane, we can simply choose $\mathbf{z} = \hat{\mathbf{y}}$. In space, the vector $\mathbf{z} = \mathbf{y}\times(\mathbf{x}\times\mathbf{y})$ will do the job.

In both cases, we can write: $\mathbf{x} = \mathbf{p} + \mathbf{q} = a\mathbf{y} + b\mathbf{z}$ and try to determine b. This can be done by performing the dot product with $\mathbf{z}$, which yields: $\mathbf{x}\cdot\mathbf{z} = b(\mathbf{z}\cdot\mathbf{z})$ and hence $b = \frac{(\mathbf{x}\cdot\mathbf{z})}{\mathbf{z}\cdot\mathbf{z}}$.

In the plane case, substitute $\mathbf{z} = \hat{\mathbf{y}}$ and use $\mathbf{x}\cdot\hat{\mathbf{y}} = [\mathbf{y},\mathbf{x}]$ from Sect. 1.2.6 to obtain the results stated in Cor. 1.27. Substituting $\mathbf{z} = \mathbf{y}\times(\mathbf{x}\times\mathbf{y})$ in the "space case", we get for the numerator:

$$\mathbf{x}\cdot\mathbf{z} = \mathbf{x}\cdot(\mathbf{y}\times(\mathbf{x}\times\mathbf{y})) = [\mathbf{x},\mathbf{y},\mathbf{x}\times\mathbf{y}] = (\mathbf{x}\times\mathbf{y})\cdot(\mathbf{x}\times\mathbf{y}) = |(\mathbf{x}\times\mathbf{y})|^2.$$

For the denominator, we arrive at $\mathbf{z}\cdot\mathbf{z} = |\mathbf{z}|^2 = |\mathbf{y}\times(\mathbf{x}\times\mathbf{y})|^2 = |\mathbf{y}|^2|\mathbf{x}\times\mathbf{y}|^2$, since $\mathbf{y}\perp\mathbf{x}\times\mathbf{y}$. Altogether, we get $b = \frac{|(\mathbf{x}\times\mathbf{y})|^2}{|\mathbf{y}|^2|\mathbf{x}\times\mathbf{y}|^2} = \frac{1}{|\mathbf{y}|^2}$. Hence $\mathbf{q} = b\mathbf{z} = \mathbf{y}\times\frac{(\mathbf{x}\times\mathbf{y})}{\mathbf{y}\cdot\mathbf{y}}$. □

EXAMPLE 1.29.
Let $\mathbf{y} = [2,1,2]$, $\mathbf{x} = [1,-1,1]$. Using Prop. 1.26, we can calculate the orthogonal projection of $\mathbf{x}$ on $sp(\mathbf{y})$ as follows:
$\mathbf{p} = \left(\frac{\mathbf{x}\cdot\mathbf{y}}{\mathbf{y}\cdot\mathbf{y}}\right)\mathbf{y} = \left(\frac{[1,-1,1]\cdot[2,1,2]}{[2,1,2]\cdot[2,1,2]}\right)[2,1,2] = \frac{3}{9}[2,1,2] = \frac{1}{3}[2,1,2]$. The difference vector $\mathbf{q} = \mathbf{x} - \mathbf{p} = [\frac{1}{3}, -\frac{4}{3}, \frac{1}{3}]$ is orthogonal to $\mathbf{y}$, since $\mathbf{q}\cdot\mathbf{y} = [\frac{1}{3}, -\frac{4}{3}, \frac{1}{3}]\cdot[2,1,2] = 0$.

3.1.2. *Projection of a point on a line.* Using orthogonal projections of *vectors* as above, we can now describe the orthogonal projection of a *point* R on a *line* l in the Euclidean plane or in Euclidean space: Let the line l through P have a parallel vector $\mathbf{y}$ (cf. (1.9)). Then, the *orthogonal projection* of R to the line l is *the* point $R_l \in l$ such that $\overrightarrow{R_lR}$ is *orthogonal* to the parallel vector $\mathbf{y}$.

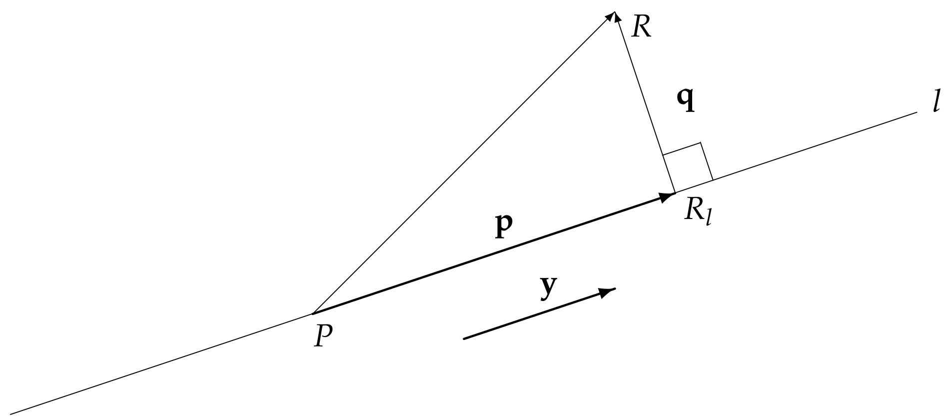


FIGURE 13. Projection of the point R on the line l

With P and R as above, we get a decomposition of the vector $\overrightarrow{PR} = \overrightarrow{PR_l} + \overrightarrow{R_lR}$ with $\mathbf{p} = \overrightarrow{PR_l}$ parallel to $\mathbf{y}$ and $\mathbf{q} = \overrightarrow{R_lR}$ orthogonal to $\mathbf{y}$. In other words: The vector $\mathbf{p} = \overrightarrow{PR_l}$ is the orthogonal projection of $\mathbf{x} = \overrightarrow{PR}$ on $\mathbf{y}$, and the point R_l is determined by the vector $\overrightarrow{OR_l} = \overrightarrow{OP} + \overrightarrow{PR_l}$.

EXAMPLE 1.30.
Let two points $P,\ R \in \mathbf{E}^3$ be given by their coordinates $P[2,1,0]$, $R[3,0,1]$; moreover the vector $\mathbf{y} = [2,1,2]$ from Ex. 1.29. We wish to calculate the orthogonal projection of R on the line $l = \{Q \in \mathbf{E}^3 | \overrightarrow{OQ} = \overrightarrow{OP} + t\mathbf{y},\ t \in \mathbf{R}\}$, cf. Fig. 13 Since $\mathbf{x} = \overrightarrow{PR} = [1,-1,1]$ as in Ex. 1.29, we get from the calculation there: $\mathbf{p} = \overrightarrow{PR_l} = \frac{1}{3}[2,1,2]$. Thus, the coordinates of the point R_l can be calculated as follows:
$\overrightarrow{OR_l} = \overrightarrow{OP} + \overrightarrow{PR_l} = [2,1,0] + \frac{1}{3}[2,1,2] = \frac{1}{3}[8,4,2]$.

3.1.3. *Projection of a point on a plane.* Let now $\alpha \subset \mathbf{E}^3$ denote a plane in Euclidean space. Remember that a vector $\mathbf{n} \neq \mathbf{0}$ is called a *normal vector* to α, if it is orthogonal to every parallel vector $\overrightarrow{PQ}$ with $P, Q \in \alpha$ (cf. Section 2.2.2).

To *define* the orthogonal projection $R_\alpha \in \alpha$ of R on α, we require that the vector $\overrightarrow{RR_\alpha}$ is a *normal vector* to α. This time, the calculation of Rs projection uses a *normal* vector $\mathbf{n}$ to α: Choose an arbitrary point $P \in \alpha$. The vector equation $\overrightarrow{PR} = \overrightarrow{PR_\alpha} + \overrightarrow{R_\alpha R} = \mathbf{q} + \mathbf{p}$, cf. Fig. 14, decomposes $\overrightarrow{PR}$ into a vector $\mathbf{q} = \overrightarrow{PR_\alpha}$ *parallel* to α and a vector $\mathbf{p} = \overrightarrow{R_\alpha R}$

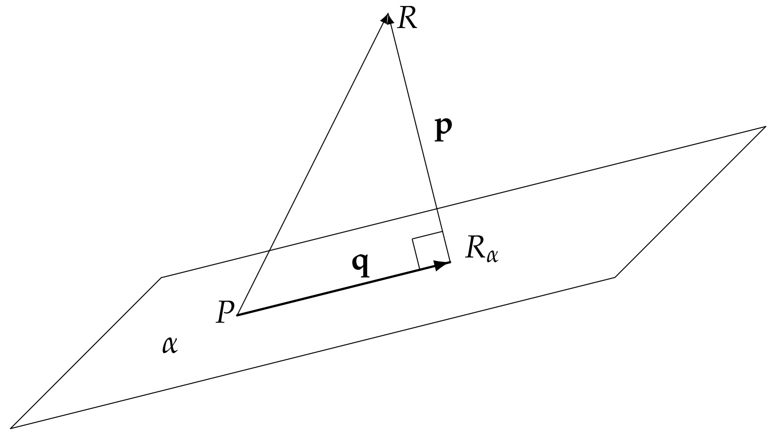


FIGURE 14. Projection of the point R on the plane α

normal to α and thus *parallel* to $\mathbf{n}$. In particular, we may calculate $\mathbf{p}$ as the orthogonal projection of $\overrightarrow{PR}$ to any normal vector $\mathbf{n}$ to α by (1.28):

$$\mathbf{p} = \frac{\overrightarrow{PR} \cdot \mathbf{n}}{\mathbf{n} \cdot \mathbf{n}} \mathbf{n}.$$

Finally, we find coordinates using $\overrightarrow{OR_\alpha} = \overrightarrow{OR} - \overrightarrow{R_\alpha R} = \overrightarrow{OR} - \mathbf{p}$.

EXAMPLE 1.31.
Let $P[1,2,1]$, $R[0,0,1] \in \mathbf{E}^3$, $\mathbf{x} = [-2,0,2]$, $\mathbf{y} = [1,-1,3] \in \mathbf{R}^3$. A plane $\alpha \subset \mathbf{E}^3$ is given by the parametrization $\alpha = \{Q \in \mathbf{E}^3 | \overrightarrow{OQ} = \overrightarrow{OP} + s\mathbf{x} + t\mathbf{y}, \quad s,t \in \mathbf{R}\}$. The vector $\mathbf{x} \times \mathbf{y} = [2,8,2]$ is normal to $sp(\mathbf{x},\mathbf{y})$, and hence $\mathbf{n} = [1,4,1]$ is normal to α as well. The orthogonal projection $\mathbf{p}$ of $\overrightarrow{PR} = [-1,-2,0]$ on $\mathbf{n}$ is calculated using Prop. 1.26: $\mathbf{p} = \overrightarrow{R_\alpha R} = \left(\frac{[-1,-2,0]\cdot[1,4,1]}{[1,4,1]\cdot[1,4,1]}\right)[1,4,1] = \frac{-9}{18}[1,4,1] = [-0.5,-2,-0.5]$. The coordinates of R_α are thus found as follows: $\overrightarrow{OR_\alpha} = \overrightarrow{OR} - \overrightarrow{R_\alpha R} = [0,0,1] + [0.5,2,0.5] = [0.5,2,1.5]$.

3.1.4. *Projection of a line on a plane.* Let $l, \alpha \subset \mathbf{E}^3$ denote a line, resp. a plane in space. The orthogonal projection $l_\alpha = \{P_\alpha | P \in l\}$ consists of the orthogonal projections of all points on l.

EXAMPLE 1.32.
With R and α as in Ex. 1.31, and $\mathbf{z} = [0,9,0]$, let l be the line through R with $\mathbf{z}$ as a parallel vector (with parametrization $\overrightarrow{OQ} = \overrightarrow{OR} + t\mathbf{z}$, $t \in \mathbf{R}$). We calculate the projection $\mathbf{p}_1$ of $\mathbf{z}$ on $\mathbf{n}$ using Prop. 1.26: $\mathbf{p}_1 = \left(\frac{[0,9,0]\cdot[1,4,1]}{[1,4,1]\cdot[1,4,1]}\right)[1,4,1] = [2,8,2]$. Hence, the projection of $\mathbf{z}$ on the parallel plane $sp(\mathbf{x},\mathbf{y})$ is $\mathbf{q}_1 = \mathbf{z} - \mathbf{p}_1 = [-2,1,-2]$, and the projection l_α on l has parametrization $\overrightarrow{OQ} = \overrightarrow{OR_\alpha} + t\mathbf{q}_1$, $t \in \mathbf{R}$, i.e., $\overrightarrow{OQ} = [0.5,2,1.5] + t[-2,1,-2]$, $t \in \mathbf{R}$.

3.2. Distances and Angles.

3.2.1. *Distances.* The *distance* $d(P,Q)$ between two *points* $P,Q \in \mathbf{E}^i$ is defined as the length $|\overrightarrow{PQ}|$ (cf. (1.5)) of the vector $\overrightarrow{PQ} \in \mathbf{R}^i$ from P to Q.

What is the distance of a point R from a line l - or a plane α? It is defined as the *shortest* distance $d(R,Q)$ from R to any point Q in l (or α). Likewise, the distance $d(l_1,l_2)$ between two lines (l_1,l_2) is the *shortest* distance $d(P_1,P_2)$ with $P_1 \in l_1$ and $P_2 \in l_2$. How can we calculate those shortest distances?

An application of *Pythagoras theorem* to the situation in Fig. 13 shows easily, that the orthogonal projection R_l is the point on the line l closest to R: For a general point P on l, we have: $d(P,R)^2 = d(P,R_l)^2 + d(R_l,R)^2 \geq d(R_l,R)^2$, cf. Fig. 13. In the same way, you find out, that $d(R,\alpha) = d(R,R_\alpha)$.

EXAMPLE 1.33.
In Ex. 1.31, we calculated the projection R_α of R to the plane α. Hence, the distance between R and α is:
$d(R,\alpha) = d(R,R_\alpha) = |\overrightarrow{RR_\alpha}| = |[0.5,2,0.5]| = \sqrt{\frac{1}{4}+4+\frac{1}{4}} = \sqrt{\frac{9}{2}} = \frac{3}{2}\sqrt{2} \sim 2.121.$

The determination of the distance between *two lines* l_1 and l_2 is more subtle, and we have to look at two cases:

(1) The lines l_1 and l_2 are *parallel*, i.e., they have a common parallel vector $\mathbf{y}$ and they are contained in a common plane $\alpha \subset \mathbf{E}^3$, cf. the left side of Fig. 15: Let $\mathbf{n} \in \mathbf{R}^3$ denote a vector perpendicular to $\mathbf{y}$ such that $\mathbf{y}$ and $\mathbf{n}$ span a plane parallel to α. Let $Q_1 \in l_1$ and $Q_2 \in l_2$ denote arbitrary points. The vector $\overrightarrow{Q_1Q_2}$ is parallel to α. Hence, there are $s,t \in \mathbf{R}$ such that
$\overrightarrow{Q_1Q_2} = s\mathbf{y} + t\mathbf{n} = \overrightarrow{Q_1P_1} + \overrightarrow{P_1Q_2}$ with $P_1 \in l_1$ and $\overrightarrow{P_1Q_2}$ parallel to $\mathbf{n}$ and thus perpendicular to both l_1 and l_2. The points Q_1, P_1 and Q_2 form a right-angled traingle (cf. Fig. 15); hence $|\overrightarrow{P_1Q_2}| \leq |\overrightarrow{Q_1Q_2}|$ for all $Q_1 \in l_1$, again by Pythagoras theorem. It is not difficult to see, that $|\overrightarrow{P_1Q_2}|$ is independent of the choice of the points Q_1 and Q_2 above, and therefore $d(l_1,l_2) = |\overrightarrow{P_1Q_2}| = \mathbf{q}$ is the minimal distance.

Since $\mathbf{p} = \overrightarrow{Q_1P_1}$ is the orthogonal projection of $\mathbf{x} = \overrightarrow{P_1Q_2}$ on $\mathbf{y}$, we can use (1.30) to calculate:
$$|\mathbf{q}| = |\overrightarrow{P_1Q_2}| = \frac{|\overrightarrow{Q_1Q_2} \times \mathbf{y}|}{|\mathbf{y}|}.$$

(2) The lines l_1 and l_2 are either *intersecting* or *skew*, i.e, their parallel vectors $\mathbf{x}$ and $\mathbf{y}$ are *not* parallel; cf. Sect. 2.3.2 and the right side of Fig. 15. The vectors $\mathbf{x}$ and $\mathbf{y}$ have then $\mathbf{n} = \mathbf{x} \times \mathbf{y}$ as a common normal. Choose two arbitrary points $Q_1 \in l_1, Q_2 \in l_2$. Since $\mathbf{x}, \mathbf{n},$ and $\mathbf{y}$ are linearly independent, one may decompose $\overrightarrow{Q_1Q_2} = r\mathbf{x} + t\mathbf{n} + s\mathbf{y} = \overrightarrow{Q_1P_1} + \overrightarrow{P_1P_2} + \overrightarrow{P_2Q_2}$ with $P_1 \in l_1$ and $P_2 \in l_2$ the *unique*

points with $\overrightarrow{P_1P_2}$ orthogonal on both l_1 and l_2. In particular, $\mathbf{p} = \overrightarrow{P_1P_2}$ is the *orthogonal projection* of $\overrightarrow{Q_1Q_2}$ on the normal vector $\mathbf{n}$. Moreover, Pythagoras theorem (cf. Fig. 15) shows, that

$$d(Q_1, Q_2)^2 = |\overrightarrow{Q_1Q_2}|^2 = |\overrightarrow{P_1P_2}|^2 + |\overrightarrow{Q_1P_1} + \overrightarrow{P_2Q_2}|^2 \geq |\overrightarrow{P_1P_2}|^2 = d(P_1, P_2)^2$$

for every choice $Q_1 \in l_1$ and $Q_2 \in l_2$. As a consequence, $\overrightarrow{P_1P_2}$ is the *shortest* vector amongst those from l_1 to l_2. Using Prop. 1.26, we can calculate

$$\overrightarrow{P_1P_2} = \left(\frac{\overrightarrow{Q_1Q_2} \cdot (\mathbf{x} \times \mathbf{y})}{|\mathbf{x} \times \mathbf{y}|^2} \right) (\mathbf{x} \times \mathbf{y}).$$

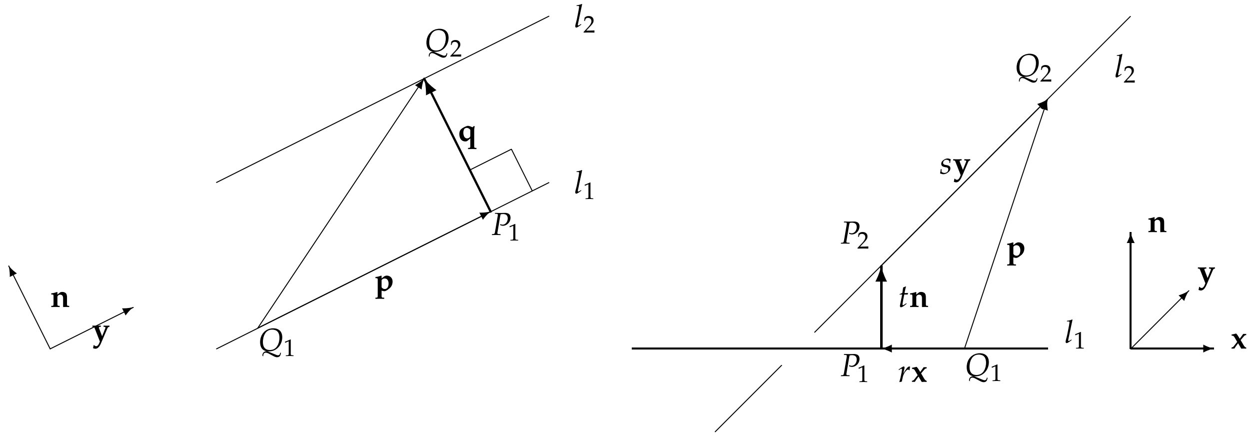


FIGURE 15. The distance between two lines in space

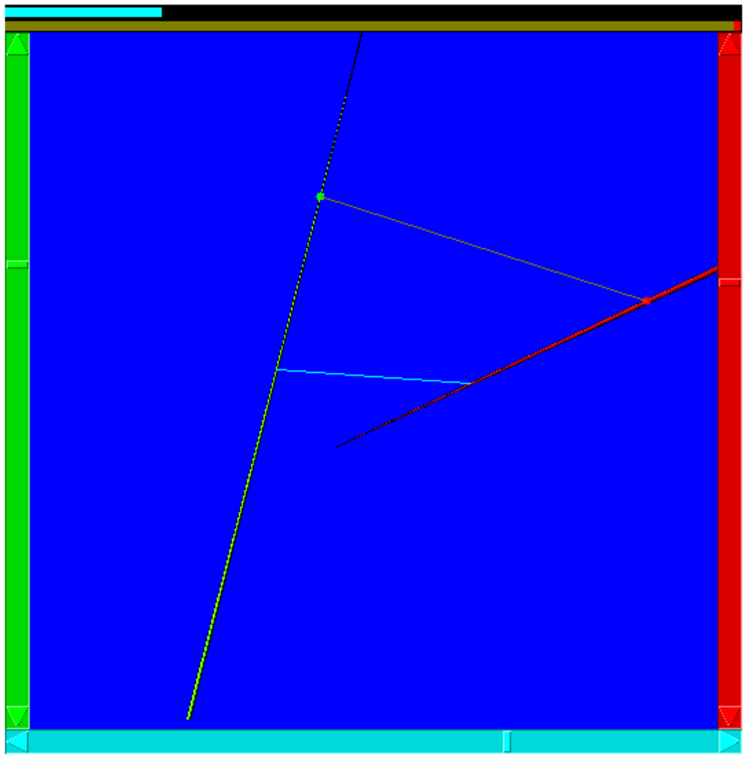
Summing up, we obtain

PROPOSITION 1.34. *Let l_1, l_2 denote lines with parallel vectors $\mathbf{x}$, resp. $\mathbf{y}$. Then, there are points $P_1 \in l_1, P_2 \in l_2$ with $\overrightarrow{P_1P_2}$ orthogonal on both l_1 and l_2, and $d(l_1, l_2) = |\overrightarrow{P_1P_2}|$. Moreover, if $Q_1 \in l_1$ and $Q_2 \in l_2$ are chosen arbitrarily, then*

$$d(l_1, l_2) = \frac{|\overrightarrow{Q_1Q_2} \times \mathbf{y}|}{|\mathbf{y}|} \text{ if } \mathbf{x} \text{ and } \mathbf{y} \text{ are parallel;}$$

$$d(l_1, l_2) = \frac{|\overrightarrow{Q_1Q_2} \cdot (\mathbf{x} \times \mathbf{y})|}{|\mathbf{x} \times \mathbf{y}|} = \frac{|[\overrightarrow{Q_1Q_2}, \mathbf{x}, \mathbf{y}]|}{|\mathbf{x} \times \mathbf{y}|} \text{ if } \mathbf{x} \text{ and } \mathbf{y} \text{ are not parallel.}$$

Here are two skew lines:



Rotate the two line segments with the mouse. Change the distance between them using the lower scrollbar. Push the red points along them using the scrollbars to the left and right. The cyan line shows the shortest distance. The umber line shows the distance between the two red points.

FIGURE 16. Distance between two skew lines

EXAMPLE 1.35.
Let

$$\begin{aligned} l_1 : \overrightarrow{OP} &= [1,2,1] + s[1,1,1], \quad s \in \mathbf{R} \\ l_2 : \overrightarrow{OQ} &= [-1,0,1] + t[0,2,-1], \quad t \in \mathbf{R}. \end{aligned}$$

With $\mathbf{x} = [1,1,1]$, $\mathbf{y} = [0,2,-1]$, $Q_1 : [1,2,1]$ and $Q_2 : [-1,0,1]$, we obtain from Prop. 1.34: $\overrightarrow{Q_1Q_2} = [-2,-2,0]$, $\mathbf{n} = \mathbf{x} \times \mathbf{y} = [-3,1,2]$, $\overrightarrow{Q_1Q_2} \cdot \mathbf{n} = [-2,-2,0] \cdot [-3,1,2] = 4$, and $d(l_1, l_2) = \frac{4}{\sqrt{14}} = \frac{2}{7}\sqrt{14} \sim 1.069$.

It is not difficult either to find the points $P_1 \in l_1, P_2 \in l_2$ *realizing* that shortest distance. We have to find $r, s \in \mathbf{R}$ with $r\mathbf{x} + s\mathbf{y} = \overrightarrow{Q_1P_1} + \overrightarrow{P_2Q_2} = \overrightarrow{Q_1Q_2} - \overrightarrow{P_1P_2}$.

EXAMPLE 1.36.
Continuing Ex. 1.35, $\overrightarrow{P_1P_2} = \left(\frac{\overrightarrow{Q_1Q_2}\cdot(\mathbf{x}\times\mathbf{y})}{|\mathbf{x}\times\mathbf{y}|^2}\right)(\mathbf{x}\times\mathbf{y}) = \frac{4}{14}[-3,1,2] = \frac{1}{7}[-6,2,4]$. Hence, we have to solve the equation
$r\mathbf{x}+s\mathbf{y} = \overrightarrow{Q_1Q_2} - \overrightarrow{P_1P_2} = [-2,-2,0] - \frac{1}{7}[-6,2,4] = \frac{-1}{7}[8,16,4]$, with $r = \frac{-8}{7}$, and $s = \frac{-4}{7}$ as a result. Finally,
$\overrightarrow{OP_1} = \overrightarrow{OQ_1} + \overrightarrow{Q_1P_1} = \overrightarrow{OQ_1} + r\mathbf{x} = [1,2,1] - \frac{8}{7}[1,1,1] = \frac{1}{7}[-1,6,-1]$, and
$\overrightarrow{OP_2} = \overrightarrow{OQ_2} - \overrightarrow{P_2Q_2} = \overrightarrow{OQ} - s\mathbf{y} = [-1,0,1] + \frac{4}{7}[0,2,-1] = \frac{1}{7}[-7,8,3]$.

3.2.2. *Angles.* We are already familiar with the angle between two vectors $\mathbf{x}$ and $\mathbf{y}$ in the plane or in space, cf. (1.6). What are the angles between

(1) two (not necessarily intersecting) lines;
(2) a line and a plane in space;
(3) two planes in space?

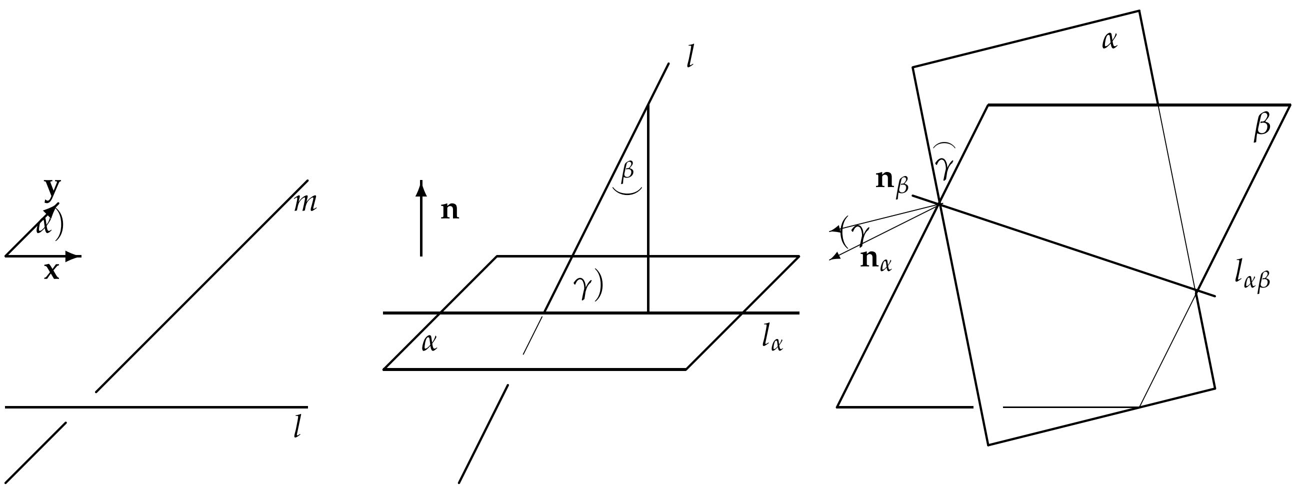


FIGURE 17. Angles

(1) Geometrically, the angle between two lines l and $m \subset \mathbf{E}^3$ is defined as the angle between l and a line $m_1 \subset \mathbf{E}^3$ parallel to m and *interesecting* l. If $\mathbf{x}$ and $\mathbf{y}$ denote parallel vectors for l and m, this angle is just the angle between those parallel vectors.

EXAMPLE 1.37.
The angle α between the lines l_1 and l_2 from Ex. 1.35 is calculated as follows: $\cos\alpha = \frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}||\mathbf{y}|} = \frac{1}{\sqrt{3}\sqrt{5}} = \frac{1}{15}\sqrt{15}$. This corresponds to an angle $\alpha \sim 1.504$, or $\alpha \sim 86.18^0$.

(2) The angle γ between a line l and a plane α in space is defined as the acute angle between l and its projection l_α on α. Let $\mathbf{n} \in \mathbf{R}^3$ denote a normal vector to the plane α, and let β denote the acute angle between l and $\mathbf{n}$ – that may be calculated as in 1. Fig. 17 shows, that $\gamma + \beta = \frac{\pi}{2}$, or $\gamma = \frac{\pi}{2} - \beta$.

EXAMPLE 1.38.
Let l be given by the parametrization $l : \overrightarrow{OQ} = [1,2,1] + s\mathbf{x}$, $s \in \mathbf{R}$, with $\mathbf{x} = [1,1,1]$, and α by the equation $x + 2y + 3z = 4$. Then $\mathbf{n} = [1,2,3]$ is a normal vector to α, and $\cos\beta = \frac{\mathbf{x}\cdot\mathbf{n}}{|\mathbf{x}||\mathbf{n}|} = \frac{6}{\sqrt{3}\sqrt{14}} = \frac{\sqrt{42}}{7} \sim 0.926$, corresponding to an angle $\beta \sim 0.388$; hence $\gamma \sim 1.183$, or $\gamma \sim 67.79^0$.

(3) Let α and β denote non-parallel planes in space, intersecting each other in a line $l_{\alpha\beta} = \alpha \cap \beta$. The angle γ between α and β is defined as an angle in a plane δ perpendicular to $l_{\alpha\beta}$, i.e., the angle between the *lines* $\alpha \cap \delta$ and $\beta \cap \delta$. For calculations, it is easier to use the following recipe: Let $\mathbf{n}_\alpha$ and $\mathbf{n}_\beta$ denote normal vectors to α, resp. β. The angle γ is then the same as the angle between the vectors $\mathbf{n}_\alpha$ and $\mathbf{n}_\beta$.

EXAMPLE 1.39.
Let the planes α and β be given by the equations $2x + y + z = 3$, resp. $x + y = 2$. They have $\mathbf{n}_\alpha = [2,1,1]$, resp. $\mathbf{n}_\beta = [1,1,0]$ as normal vectors. The angle γ between α and β is then determined by $\cos\gamma = \frac{\mathbf{n}_\alpha\cdot\mathbf{n}_\beta}{|\mathbf{n}_\alpha||\mathbf{n}_\beta|} = \frac{[2,1,1]\cdot[1,1,0]}{|[2,1,1]||[1,1,0]|} = \frac{3}{\sqrt{6}\sqrt{2}} = \frac{\sqrt{3}}{2}$, and hence $\gamma = \frac{\pi}{6}$ corresponding to an angle of 30^0.

3.3. Examples of Applications within Mechanics.

EXAMPLE 1.40.
A box weighing 100 kg is placed on an incline of angle 50°. What is the magnitude of the component of the force of gravity along the incline? What is the magnitude of the component normal to the incline? (The gravitational acceleration is $g = 9.81m/s^2$.)
Solution: The total force of gravity is 981N, and the direction is vertical. Hence the component along the incline is 981N$\cdot \sin 50^\circ = 751$N and the component normal to the incline is 981N$\cdot \cos 50^\circ = 631$N.

EXAMPLE 1.41.
A force $\mathbf{F}$ with P as its point of application has a moment $\mathbf{M}_O = \overrightarrow{OP} \times \mathbf{F}$ about the point O. The line of action of $\mathbf{F}$ is the line defined by $\mathbf{F}$ and P. Let Q be a point on the line of action and let $\mathbf{F}$ attack at Q; the moment about O is then

$$\mathbf{M}_O = \overrightarrow{OQ} \times \mathbf{F} = (\overrightarrow{OP} + \overrightarrow{PQ}) \times \mathbf{F} = (\overrightarrow{OP} + t\mathbf{F}) \times \mathbf{F} = \overrightarrow{OP} \times \mathbf{F} + t\mathbf{F} \times \mathbf{F} = \overrightarrow{OP} \times \mathbf{F}.$$

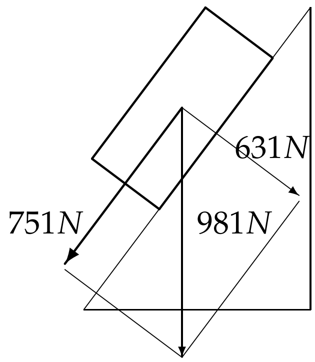


FIGURE 18. A box on an incline

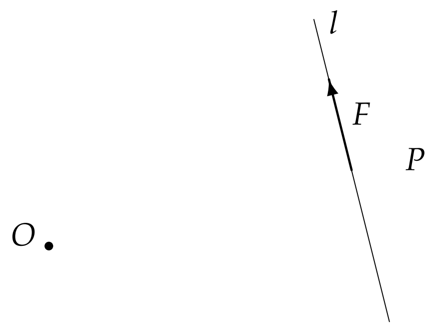


FIGURE 19. The line of action

Hence, the moment depends on the line of action and *not* on the specific point P. Note that O is not necessarily the origin in this calculation.

It follows that if O is on the line of action, the moment of $\mathbf{F}$ about O is zero: we just assume that F attacks at O!

EXAMPLE 1.42.
A rectangular parallelepid $ABCDEFGH$ with edges $AB = 4$m, $AD = 5$m, and $AE = 3$m is acted on by a force $\mathbf{F}$ of magnitude 20N along the diagonal EC. What is the moment about A and what are the components of $\mathbf{F}$ along the edges of the parallelepid?

Solution: Place a coordinate system with A as the origin, x-axis along AB, y-axis along AD and z-axis along AE. The coordinates of $\overrightarrow{EC}$ are thus $[4,5,-3]$ and $|EC| = \sqrt{4^2+5^2+3^2} = \sqrt{50}$. A unit vector in the direction of $\overrightarrow{EC}$ is $\frac{\sqrt{2}}{10}[4,5,-3]$. Since the magnitude of $\mathbf{F}$ is 20N, we have $\mathbf{F} = [F_x, F_y, F_z] = 2\sqrt{2}[4,5,-3]$N.

The moment about A is

$$\mathbf{M} = \overrightarrow{AC} \times \mathbf{F} = \sqrt{2}[-30, 24, 0]Nm.$$

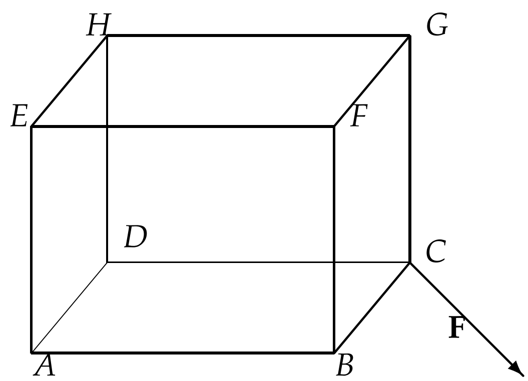


FIGURE 20. The parallelepipedon ABCDEFGH

The magnitude of **M** is $|M| = 6\sqrt{82} \sim 54.33$Nm. Since $\mathbf{F} = 2\sqrt{2}[4, 5, -3]$N, the magnitude of the components along the axis are read off to be $F_x = 8\sqrt{2}$N, $F_y = 10\sqrt{2}$N og $F_z = -6\sqrt{2}$N.

EXAMPLE 1.43.
The equilibrium conditions for a body are the conditions that ensure, that the body is either fixed or moves with constant velocity. In calculations, one fixes a point A. The equilibrium conditions then mean, that the sum of all forces acting on the body, and the sum of the moments of these forces about A are both trivial. We will show, that this requirement is independent of the choice of the point A.

Assume, that there are k forces $\mathbf{F}_1, \mathbf{F}_2, \dots \mathbf{F}_k$ acting on the body and that P_i is a point on the line of action of $\mathbf{F}_i$, i = 1,..,k. The equilibrium conditions yield: $\sum_{i=1}^k \mathbf{F}_i = 0$ and $\sum_{i=1}^k \overrightarrow{AP_i} \times \mathbf{F}_i = 0$.

Assume that these conditions are fulfilled and let B be another point. The equilibrium conditions for B are then: $\sum_{i=1}^k \mathbf{F}_i = 0$ and $\sum_{i=1}^k \overrightarrow{BP_i} \times \mathbf{F}_i = 0$. The first condition is the same as above and thus satisfied. For the latter condition, we calculate:

$$\sum_{i=1}^k \overrightarrow{BP_i} \times \mathbf{F}_i = \sum_{i=1}^k (\overrightarrow{BA} + \overrightarrow{AP_i}) \times \mathbf{F}_i = \sum_{i=1}^k \overrightarrow{BA} \times \mathbf{F}_i + \sum_{i=1}^k \overrightarrow{AP_i} \times \mathbf{F}_i = \overrightarrow{BA} \times \sum_{i=1}^k \mathbf{F}_i = 0.$$

A smart choice of the point may simplify calculations considerably; remember that the moment about a point on the line of action is 0. Hence a point on the line of action of as many of the forces involved as possible is usually a smart choice.

4. Change of Coordinate Systems

In many practical situations, it turns out to be useful to adapt the description of geometrical phenomena by carefully choosing a coordinate system: Imagine we want to describe a crater on the moon. Of course, we would choose a coordinate system with center in that crater rather than in our home town. In the same way, one may have to turn an object (or the camera – i.e., the coordinate system) around in order to see all the important features of it. If an object is represented on a computer screen and turned around there, the computer program has to perform calculations that correspond to continuous changes of coordinate systems.

4.1. Rotation around the base point. Let us look at the simplest case first. Suppose we are given two systems $\{\mathbf{i},\mathbf{j}\}$ and $\{\mathbf{i}',\mathbf{j}'\}$ of plane coordinate systems – consisting of pairwise *orthogonal unit* vectors, satisfying

$$|\mathbf{i}| = |\mathbf{j}| = |\mathbf{i}'| = |\mathbf{j}'| = 1;\ \mathbf{j} = \hat{\mathbf{i}},\ \mathbf{j}' = \hat{\mathbf{i}'}.$$

How many *bases* $\{\mathbf{i}',\mathbf{j}'\}$ with these properties are there? Well, $\mathbf{i}'$ can be obtained by rotating $\mathbf{i}$ (counterclockwise) through an angle v. But then, $\mathbf{j}' = \hat{\mathbf{i}'}$ is obtained by rotating $\mathbf{i}'$ by $\frac{\pi}{2}$, resp. $\mathbf{i}$ by $\frac{\pi}{2}+v$, or by rotating $\mathbf{j} = \hat{\mathbf{i}}$ by v, cf. Fig. 21. Hence, the whole coordinate system is rotated through an angle v. More explicitly we get the following formulae expressing *change of basis*:

$$\begin{aligned} \mathbf{i}' &= \cos v\mathbf{i} + \sin v\mathbf{j}, \\ \mathbf{j}' = \hat{\mathbf{i}'} &= -\sin v\mathbf{i} + \cos v\mathbf{j}. \end{aligned} \tag{1.31}$$

What happens to the coordinates of a point $P \in \mathbf{E}^2$ when expressed in the two (old and new) coordinate systems? Assume that P has coordinates $[x,y]$ with respect to $\{\mathbf{i},\mathbf{j}\}$ and $[x',y']$ with respect to $\{\mathbf{i}',\mathbf{j}'\}$.

PROPOSITION 1.44. (*Change of* coordinates:)
P's coordinate sets are related as follows:

$$\begin{aligned} x &= x'\cos v - y'\sin v \\ y &= x'\sin v + y'\cos v, \end{aligned} \tag{1.32}$$

$$\begin{aligned} x' &= x\cos v + y\sin v \\ y' &= -x\sin v + y\cos v \end{aligned} \tag{1.33}$$

PROOF:
Expressing the vector $\overrightarrow{OP}$ as a linear combination in both bases, we obtain using (1.31):

$$\begin{aligned} x\mathbf{i} + y\mathbf{j} = \overrightarrow{OP} &= x'\mathbf{i}' + y'\mathbf{j}' \\ &= x'(\cos v\mathbf{i} + \sin v\mathbf{j}) + y'(-\sin v\mathbf{i} + \cos v\mathbf{j}) \\ &= (x'\cos v - y'\sin v)\mathbf{i} + (x'\sin v + y'\cos v)\mathbf{j}, \end{aligned} \tag{1.34}$$

and (1.32) follows by comparing the coordinates with respect to the basis vectors **i** and **j**. Formula (1.33) can be shown in a similar way, observing that $\{\mathbf{i}, \mathbf{j}\}$ can be obtained by rotating $\{\mathbf{i}', \mathbf{j}'\}$ through an angle $-v$. □

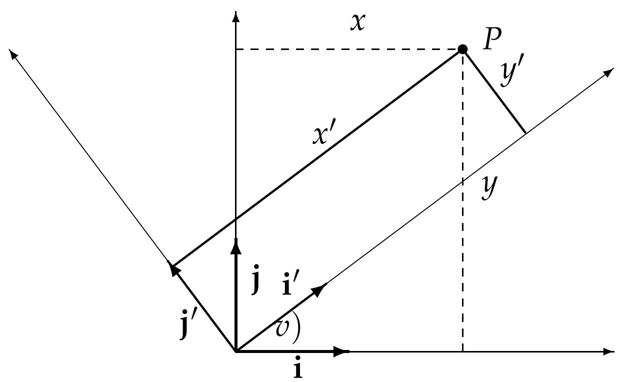


FIGURE 21. Change of basis and coordinates

REMARK 1.45.
The "new" coordinates express the *projections* of $\overrightarrow{OP}$ on the new coordinate lines through the origin with $\mathbf{i}'$, resp. $\mathbf{j}'$ as parallel vectors. It is more conceptual – and easier to

memorize – to use matrices to derive formulas (1.32) and (1.33) above; the important matrix here is

$$\mathbf{M} = \begin{bmatrix} \cos v & -\sin v \\ \sin v & \cos v \end{bmatrix} \tag{1.35}$$

Its column vectors are the vectors $\mathbf{i}'$ and $\mathbf{j}'$ expressed in the basis $\{\mathbf{i}, \mathbf{j}\}$. We make also use of **M**s transpose $\mathbf{M}^T = \begin{bmatrix} \cos v & \sin v \\ -\sin v & \cos v \end{bmatrix}$. Remark, that **M** is an *orthogonal* matrix, i.e., $\mathbf{M}\mathbf{M}^T = \mathbf{M}^T\mathbf{M} = \mathbf{I}$. (Do the calculations, using: $(\cos v)^2 + (\sin v)^2 = 1$!) In particular, **M**s inverse $\mathbf{M}^{-1}$ is just its transpose: $\mathbf{M}^{-1} = \mathbf{M}^T$.

The change of basis formulae can now be expressed as follows: Let $[\mathbf{i}, \mathbf{j}]$ and $[\mathbf{i}', \mathbf{j}']$ denote *formal* row vectors with the indicated coordinates. Their relation (1.31) is expressible by matrix multiplication as follows:

$$[\mathbf{i}', \mathbf{j}'] = [\mathbf{i}, \mathbf{j}]\mathbf{M}. \tag{1.36}$$

The relation (1.34) between *P*s coordinates can thus be expressed by matrix equations (matrix products yielding a 1×1-matrix, with entry a vector in $\mathbf{R}^2$) as follows:

$$[\mathbf{i}, \mathbf{j}] \begin{bmatrix} x \\ y \end{bmatrix} = [\mathbf{i}', \mathbf{j}'] \begin{bmatrix} x' \\ y' \end{bmatrix} = [\mathbf{i}, \mathbf{j}]\mathbf{M} \begin{bmatrix} x' \\ y' \end{bmatrix}. \tag{1.37}$$

Comparing coordinates yields (1.32): $\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{M} \begin{bmatrix} x' \\ y' \end{bmatrix}$.

Multiplication with $\mathbf{M}^{-1}$ yields (1.33): $\begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{M}^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{M}^T \begin{bmatrix} x \\ y \end{bmatrix}$.

REMARK 1.46.
The matrix transforming *coordinates* is *inverse* to the matrix transforming the *bases*.

EXAMPLE 1.47.
A counterclockwise rotation by $v = \frac{\pi}{3} = 60^0$ transforms the basis vectors $\mathbf{i}, \mathbf{j}$ into $\mathbf{i}' = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2}\sqrt{3} \end{bmatrix}$, $\mathbf{j}' = \begin{bmatrix} -\frac{1}{2}\sqrt{3} \\ \frac{1}{2} \end{bmatrix}$. Hence, rotation by 60^0 can be expressed by the orthogonal matrix $\mathbf{M} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix}$. Its inverse is its transpose:
$\mathbf{M}^{-1} = \mathbf{M}^T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix}$. If the point $P \in \mathbf{E}^2$, has coordinates $[x, y] = [1, 1]$ with respect to the basis $\{\mathbf{i}, \mathbf{j}\}$, its coordinates $[x', y']$ with respect to $\{\mathbf{i}', \mathbf{j}'\}$ can be calculated as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{M}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{M}^T \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(1+\sqrt{3}) \\ \frac{1}{2}(1-\sqrt{3}) \end{bmatrix}.$$

EXAMPLE 1.48.
Let a line l in the plane $\mathbf{E}^2$ in xy-coordinates be given by the equation $x + 2y = 3$. Which equation describes the same line l in the $x'y'$-coordinates from Ex. 1.47, i.e., with axes rotated by 60^0 counterclockwise?

$$3 = [1,2] \begin{bmatrix} x \\ y \end{bmatrix} = [1,2]\mathbf{M} \begin{bmatrix} x' \\ y' \end{bmatrix} = [1,2] \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} =$$
$$= [\frac{1}{2} + \sqrt{3}, -\frac{1}{2}\sqrt{3} + 1] \begin{bmatrix} x' \\ y' \end{bmatrix} = (\frac{1}{2} + \sqrt{3})x' + (-\frac{1}{2}\sqrt{3} + 1)y'.$$

It is easy to generalise the argument above to *three (or more) dimensions*: Let $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ and

$\{\mathbf{i}', \mathbf{j}', \mathbf{k}'\}$ denote two orthogonal coordinate systems (bases) in $\mathbf{R}^3$, satisfying the rule of thumbs: $\mathbf{k} = \mathbf{i} \times \mathbf{j}$, resp. $\mathbf{k}' = \mathbf{i}' \times \mathbf{j}'$. Again, there is a 3×3-matrix $\mathbf{M}$ (with $\mathbf{i}', \mathbf{j}', \mathbf{k}'$ as column vectors) relating the bases by the equation: $[\mathbf{i}', \mathbf{j}', \mathbf{k}'] = [\mathbf{i}, \mathbf{j}, \mathbf{k}]\mathbf{M}$. Moreover, $\mathbf{M}$ is an *orthogonal* matrix, i.e., $\mathbf{M}\mathbf{M}^T = \mathbf{M}^T\mathbf{M} = \mathbf{I}$, and again, it represents a rotation in Euclidean space. But a spatial rotation can no longer be specified by an angle alone: Moreover, you need an *axis* (an eigenvector of $\mathbf{M}$!). Nevertheless, if the point $P \in \mathbf{E}^3$

has $[x, y, z]$, resp. $[x', y', z']$ as its set of coordinates with respect to the two coordinate systems above, we get the following relations:

$$\begin{aligned} \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= \mathbf{M} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \\ \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} &= \mathbf{M}^{-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{M}^T \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned} \tag{1.38}$$

REMARK 1.49.
The formulae above are special cases of those for *general* change of basis and change of coordinates from linear algebra. Our case is particularly simple, since the matrix $\mathbf{M}$ expressing these changes is orthogonal.

4.2. Change of base point. As indicated in the introduction, it is often practical to exchange the base point O with another base point (or origin) $O' \in \mathbf{E}^i$. This amounts to a parallel translation of the coordinate system (without changing directions, i.e., parallel vectors) by the vector $\overrightarrow{OO'}$. The relations between the coordinates of a point $P \in \mathbf{E}^i$ with respect to the two coordinate systems can be read off the vector equation

$$\overrightarrow{OP} = \overrightarrow{OO'} + \overrightarrow{O'P}. \tag{1.39}$$

In particular, let $\overrightarrow{OO'} = a\mathbf{i} + b\mathbf{j} = [a, b]$ represent the coordinates of the "new" base point. Let furthermore $\overrightarrow{OP} = [x, y]$, resp. $\overrightarrow{O'P} = [x', y']$ represent Ps coordinates with respect to the old and the new base point. From (1.39), we obtain:

LEMMA 1.50. $[x', y'] = [x, y] - [a, b]$.

Finally, we want to see what happens to the coordinates of a point $P \in \mathbf{E}^i$ if the origin is translated *simultaneously* with a rotation of the coordinate directions by an angle v. With O, O' as above, the relations between the coordinates of a point $P \in \mathbf{E}^2$ are given by the equations

$$x\mathbf{i} + y\mathbf{j} = \overrightarrow{OP} = \overrightarrow{OO'} + \overrightarrow{O'P} = a\mathbf{i} + b\mathbf{j} + x'\mathbf{i}' + y'\mathbf{j}';$$
$$x'\mathbf{i}' + y'\mathbf{j}' = \overrightarrow{O'P} = \overrightarrow{O'O} + \overrightarrow{OP} = a'\mathbf{i}' + b'\mathbf{j}' + x\mathbf{i} + y\mathbf{j}$$

comparing the "old" coordinates $[x, y]$ of P with the "new" coordinates $[x', y']$. Above, $[a', b']$ are the coordinates of the "old" origin O with respect to the "new" coordinate system. Combining with the results of Prop. 1.44 and using the matrix $\mathbf{M}$ from (1.35), we obtain:

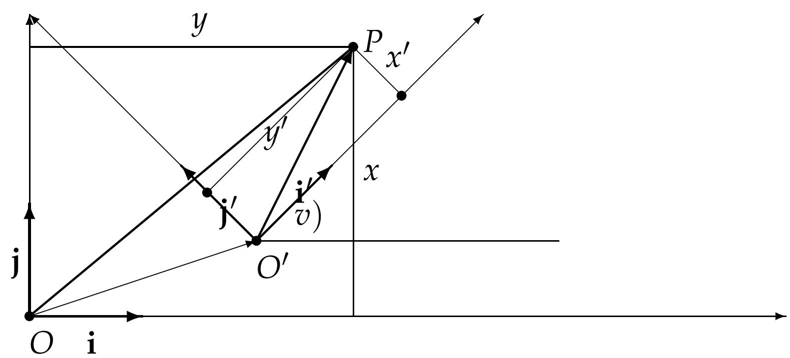


FIGURE 22. Change of coordinate system and coordinates

PROPOSITION 1.51. *Under the transition of coordinate systems* $(O, \mathbf{i}, \mathbf{j}) \mapsto (O', \mathbf{i}', \mathbf{j}')$, *the coordinates* $[x, y]$, *resp.* $[x', y']$ *of a point* $P \in \mathbf{E}^2$ *satisfy the following relations:*

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + \mathbf{M} \begin{bmatrix} x' \\ y' \end{bmatrix}; \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a' \\ b' \end{bmatrix} + \mathbf{M}^T \begin{bmatrix} x \\ y \end{bmatrix}. \tag{1.40}$$

Moreover,

$$\begin{bmatrix} a' \\ b' \end{bmatrix} = -\mathbf{M}^T \begin{bmatrix} a \\ b \end{bmatrix}. \tag{1.41}$$

PROOF:
The last asssertion (1.41) follows form the second equation in (1.40) for the coordinates of the "new" origin O', i.e., for $[x', y'] = [0, 0]$. □

REMARK 1.52.
Prop. 1.51 generalizes to shift of coordinate systems in space in the obvious way.

4.3. Application: Conic sections and quadratic equations. Change of coordinate systems and coordinates as in Prop. 1.51 can be applied to analyse a conic section C in the plane $\mathbf{E}^2$ given by a quadratic equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0. \tag{1.42}$$

We would like to find out whether the set of points $P \in \mathbf{E}^2$ with coordinates $[x, y]$ satisfying (1.42) is an ellipse, a hyperbola, a parabola, or a degenerate conic section. This can be done by introducing a particular new coordinate system $(O', \mathbf{i}', \mathbf{j}')$: It turns out, that this coordinate system - in most cases - can be obtained by a rotation through an angle v given by $\cot 2v = \frac{A-C}{B}$ and a suitable shift of base point. With respect to that system, equation (1.42) is equivalent to a simpler equation, e.g., of type

$$A'(x')^2 + C'(y')^2 + F' = 0. \tag{1.43}$$

From the latter equation, it is easy to find out what sort of conic section C represents. Instead of a general theory, we illustrate this method by two examples:

EXAMPLE 1.53.
A conic section K is defined by the equation

$$5x^2 + 2\sqrt{3}xy + 7y^2 - 32 = 0.$$

To eliminate the xy-term, we rotate the axes through an angle v in the counterclockwise direction. The angle v is the unique acute angle, which satisfies $\cot 2v = \frac{5-7}{2\sqrt{3}} = -\frac{1}{\sqrt{3}}$ and hence, $\cos v = 1/2$ and $\sin v = \frac{\sqrt{3}}{2}$, that is $v = \frac{\pi}{3}$. Using the substitutions (1.32), we arrive at the coefficients for the equation for K in the rotated coordinate system:

$$\begin{gathered} A' = 5 \cdot \frac{1}{4} + 2\sqrt{3} \cdot \frac{\sqrt{3}}{4} + 7 \cdot \frac{3}{4} = 8; \\ B' = 0; \\ C' = 5 \cdot \frac{3}{4} - 2\sqrt{3} \cdot \frac{\sqrt{3}}{4} + 7 \cdot \frac{1}{4} = 4; \\ D' = E' = 0; \\ F' = -32; \end{gathered}$$

Thus the equation defining K is

$$8x'^2 + 4y'^2 - 32 = 0 \Leftrightarrow 2x'^2 + y'^2 - 8 = 0 \Leftrightarrow \frac{x'^2}{4} + \frac{y'^2}{8} = 1.$$

Hence, K is an ellipse with major semiaxis $a = \sqrt{8} = 2\sqrt{2}$ and minor semiaxis $b = 2$. Remark that a is the semiaxis along the y'-axis, and b is the semiaxis along the x' -axis. We calculate $e = \sqrt{\frac{a^2-b^2}{a^2}} = \frac{\sqrt{2}}{2}$, $ea = 2$ and $\frac{a}{e} = 4$.

In the $[x', y']$-system the ellipse has center $[0, 0]$, vertices $[0, \pm 2\sqrt{2}]$ and $[\pm 2, 0]$, foci $[0, \pm 2]$ and directrices $y' = \pm a/e = \pm 4$.

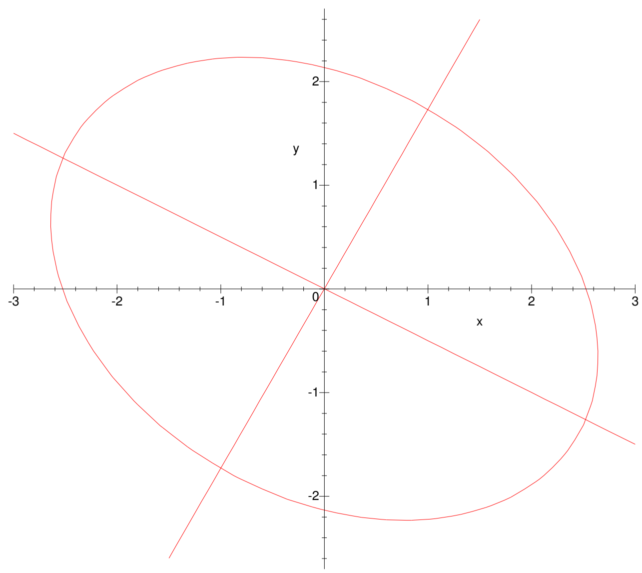


FIGURE 23. The ellipse K

To find these data in the original $[x, y]$-system, we change coordinates by the formula

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \tag{1.44}$$

Using (1.44), we find the following data in the $[x, y]$-system: center $[0, 0]$, vertices $\pm[-\sqrt{6}, \sqrt{2}]$ and $\pm[1, \sqrt{3}]$, foci $\pm[-\sqrt{3}, 1]$. Using the inverse of the matrix equation 1.44 we find the equation of the directrices:

$$\pm 4 = y' = \frac{-\sqrt{3}}{2}x + \frac{1}{2}y,$$

i.e., $y = \sqrt{3}x \pm 8$.

EXAMPLE 1.54.
We consider the conic section E with equation

$$2x^2 + 2xy + 2y^2 - 2\sqrt{2}x - 2\sqrt{2}y + 1 = 0.$$

As above, we rotate the axes through an angle v, where $\cot 2v = \frac{2-2}{2} = 0$, i.e. $v = 45°$ and $\cos v = \sin v = \frac{\sqrt{2}}{2}$. We calculate using the substitutions (1.32):

$$A' = 3,\ C' = 1,\ D' = -4,\ E' = 0 \text{ and } F' = 1.$$

The equation defining the conic section in the rotated coordinate system is then:

$$3x'^2 + y'^2 - 4x' + 1 = 0.$$

Completing the square, we get

$$3(x' - \frac{2}{3})^2 + y'^2 - \frac{1}{3} = 0.$$

To simplify this equation, we parallel translate the coordinate system $[x', y']$ to a new coordinate system $[x'', y'']$ with origin O'' in $[x', y'] = [\frac{2}{3}, 0]$. The equation for our conic section in the $[x'', y'']$-system is

$$3x''^2 + y''^2 - \frac{1}{3} = 0 \Leftrightarrow 9x''^2 + 3y''^2 = 1;$$

hence it is an ellipsis with major semiaxis $a = \frac{\sqrt{3}}{3}$ and minor semiaxis $b = \frac{1}{3}$. Again the major semiaxis is along the y''-axis. Moreover $e = \sqrt{\frac{a^2-b^2}{a^2}} = \sqrt{\frac{2}{3}}$, $ea = \frac{\sqrt{2}}{3}$ and $\frac{a}{e} = \frac{\sqrt{2}}{2}$. The directrices have equations $y'' = \pm\frac{\sqrt{2}}{2}$ and hence $y' = \pm\frac{\sqrt{2}}{2}$. To find the equations in the $[x, y]$-system we use the rotation formula and get $y' = \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y$. The equation for the directrices is then $\frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y = \pm\frac{\sqrt{2}}{2} \Leftrightarrow y = x \pm 1$.

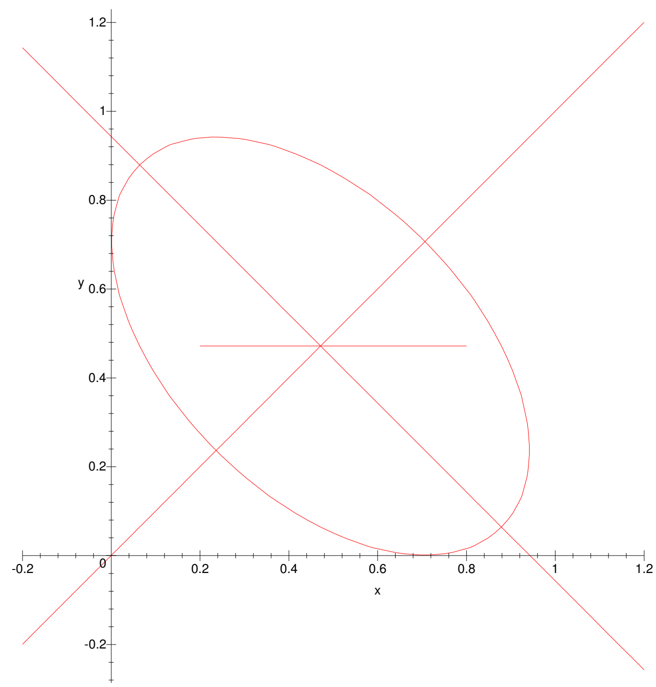


FIGURE 24. The ellipse E

The center has coordinates $[x'', y''] = [0, 0]$, thus $[x', y'] = [\frac{2}{3}, 0]$, and we use

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

to see that the center is $[x, y] = [\frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{3}]$.

The foci have coordinates $[x'', y''] = [0, \pm\frac{\sqrt{2}}{3}]$, hence $[x', y'] = [\frac{2}{3}, \pm\frac{\sqrt{2}}{3}]$ and $[x, y] = [\frac{\sqrt{2}\mp 1}{3}, \frac{\sqrt{2}\pm 1}{3}]$.

CHAPTER 2

Curves in plane and space

1. Vector functions in one variable

1.1. Definitions and elementary properties. 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 (cf. Sect. 2.1)! And for space curves, it is obvious that one has to find other means of description:

We let again $\mathbf{R}^2$, resp. $\mathbf{R}^3$ denote ordinary 2-, resp. 3-dimensional vector spaces, equipped with orthonormal bases $\{\mathbf{i}, \mathbf{j}\}$, resp. $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$.

DEFINITION 2.1. Let $a < b$ denote real numbers. A function $\mathbf{r} : [a, b] \to \mathbf{R}^3$ is called a *vector function*. A vector function can be described in coordinates as

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}, \ t \in [a, b], \tag{2.1}$$

or, in short, $\mathbf{r}(t) = [x(t), y(t), z(t)]$.

DEFINITION 2.2. A vector function $\mathbf{r} : [a, b] \to \mathbf{R}^3, \ \mathbf{r}(t) = [x(t), y(t), z(t)]$ is called *smooth* (C^∞) if the coordinate functions $x(t), y(t)$ and $z(t)$ are infinitely many times differentiable on the open interval (a, b) and continuous on $[a, b]$. For $t \in (a, b)$, its derivative $\mathbf{r}' : (a, b) \to \mathbf{R}^3$ is given by $\mathbf{r}'(t) = [x'(t), y'(t), z'(t)]$, i.e., by the derivatives of the coordinate functions.

Of course, a vector function into $\mathbf{R}^2$ (and its derivative) is defined similarly using just *two* coordinate functions.

REMARK 2.3.

(1) In this chapter, we will have to differentiate up to three times. Infinite differentiablity is imposed for simplicity of the presentation.
(2) In practice, one can often extend a smooth vector function from $[a,b]$ to some larger open interval $(a-\varepsilon, b+\varepsilon)$. In that case, continuity at a, resp. b is automatic.

PROPOSITION 2.4. *Let* $\mathbf{r}_1, \mathbf{r}_2 : [a,b] \to \mathbf{R}^i$, $i = 2$ *or* $i = 3$, *denote smooth vector functions, and let* $f : [a,b] \to \mathbf{R}$ *and* $s : [c,d] \to [a,b]$ *denote (ordinary) smooth functions. The derivatives of compound functions satisfy the following rules at every* $t \in (a,b)$:

(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*) $(\mathbf{r}_1 \circ s)'(t) = s'(t)\mathbf{r}_1'(s(t))$, $t \in (c,d)$.

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. ☐

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

PROPOSITION 2.5. (*The fundamental trick*)
Let $c \in \mathbf{R}$ *denote a constant.*

(1) *Let* $\mathbf{r}_1, \mathbf{r}_2 : [a,b] \to \mathbf{R}^i$ *denote two smooth vector functions with*
$$\mathbf{r}_1(t) \cdot \mathbf{r}_2(t) = c \text{ for every } t \in [a,b].$$
Then, $\mathbf{r}_1'(t) \cdot \mathbf{r}_2(t) + \mathbf{r}_1(t) \cdot \mathbf{r}_2'(t) = 0$ *for every* $t \in (a,b)$.
(2) *Let* $\mathbf{r} : [a,b] \to \mathbf{R}^i$ *denote a smooth vector function with* $|\mathbf{r}(t)| = c$ *for every* $t \in [a,b]$. *Then* $\mathbf{r}(t) \cdot \mathbf{r}'(t) = 0$ *for every* $t \in (a,b)$.
In particular, $\mathbf{r}'(t)$ *is* perpendicular *on* $\mathbf{r}(t)$ *for every* t.

PROOF:
Apply the product rule for vector functions (2.4.3). □

1.2. Taylor approximation.

DEFINITION 2.6. Let $\mathbf{r} : [a,b] \to \mathbf{R}^i$ denote a smooth vector function. For $k > 0$ and $t_0 \in (a,b)$, the k-th degree Taylor polynomial of $\mathbf{r}$ at t_0 is defined as

$$\mathbf{r}_{t_0}^{(k)}(t) = \mathbf{r}(t_0) + (t-t_0)\mathbf{r}'(t_0) + \frac{(t-t_0)^2}{2}\mathbf{r}''(t_0) + \cdots + \frac{(t-t_0)^k}{k!}\mathbf{r}^{(k)}(t_0)$$

$$(2.2) = \sum_{i=0}^{k} \frac{(t-t_0)^i}{i!}\mathbf{r}^{(i)}(t_0).$$

In particular,

$$\mathbf{r}_{t_0}^{(1)}(t) = \mathbf{r}(t_0) + (t-t_0)\mathbf{r}'(t_0); \tag{2.3}$$

$$\mathbf{r}_{t_0}^{(2)}(t) = \mathbf{r}(t_0) + (t-t_0)\mathbf{r}'(t_0) + \frac{(t-t_0)^2}{2}\mathbf{r}''(t_0). \tag{2.4}$$

REMARK 2.7.
The components of the vector function $\mathbf{r}_{t_0}^{(k)}$ are just the k-th degree Taylor polynomials of the components of the vector function $\mathbf{r}$ at t_0. The following result, which is proven

as in the case of ordinary functions of one variable, shows that a Taylor polynomial is a *good approximation* – in fact, the best possible approximation by a degree k polynomial vector function.

PROPOSITION 2.8. *The quotient* $\frac{\mathbf{r}(t)-\mathbf{r}_{t_0}^{(k)}(t)}{(t-t_0)^k}$ *tends to zero as t tends to t_0.*

Remark that, for $\mathbf{r}'(t_0) \neq \mathbf{0}$, the vector function $\mathbf{r}_{t_0}^{(1)}(t)$ is the parametrization of a *line* through the point P_0 with $\overrightarrow{OP_0} = \mathbf{r}(t_0)$ – the best approximating (*tangent*) line, cf. 2.2.

2. Parametrized Curves

2.1. Parametrizations. Let us get started with some examples: A parametrization of a line (cf. Def. 1.9) $\mathbf{r}(t) = \overrightarrow{OQ_t} = \overrightarrow{OP} + t\mathbf{x}, \quad t \in \mathbf{R}$, is in fact given by the *vector function* $\mathbf{r} : \mathbf{R} \to \mathbf{R}^3$ above. 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 2.9.
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.

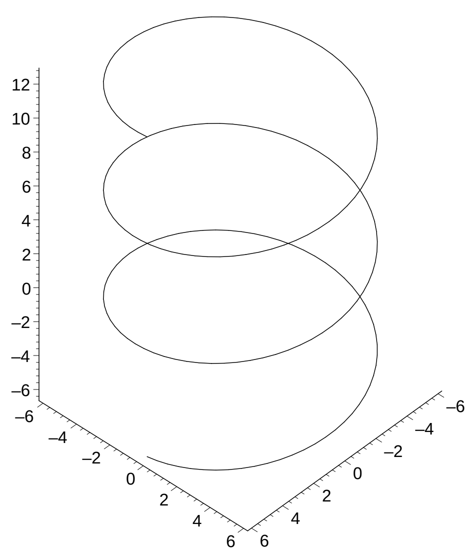


FIGURE 1. A helix

DEFINITION 2.10. A smooth vector function $\mathbf{r} : [a, b] \to \mathbf{R}^i$, $i = 2$ or 3, is called a *parametrization*[a] of the curve

$$C = \{P_t \in \mathbf{E}^i | t \in [a, b]\} \text{ given by } \overrightarrow{OP_t} = \mathbf{r}(t).$$

The parametrization $\mathbf{r}$ is called *regular* if and only if

$$\mathbf{v}(t) = \mathbf{r}'(t) \neq \mathbf{0} \text{ for every } t \in (a, b). \tag{2.5}$$

[a] illustration: below and next page

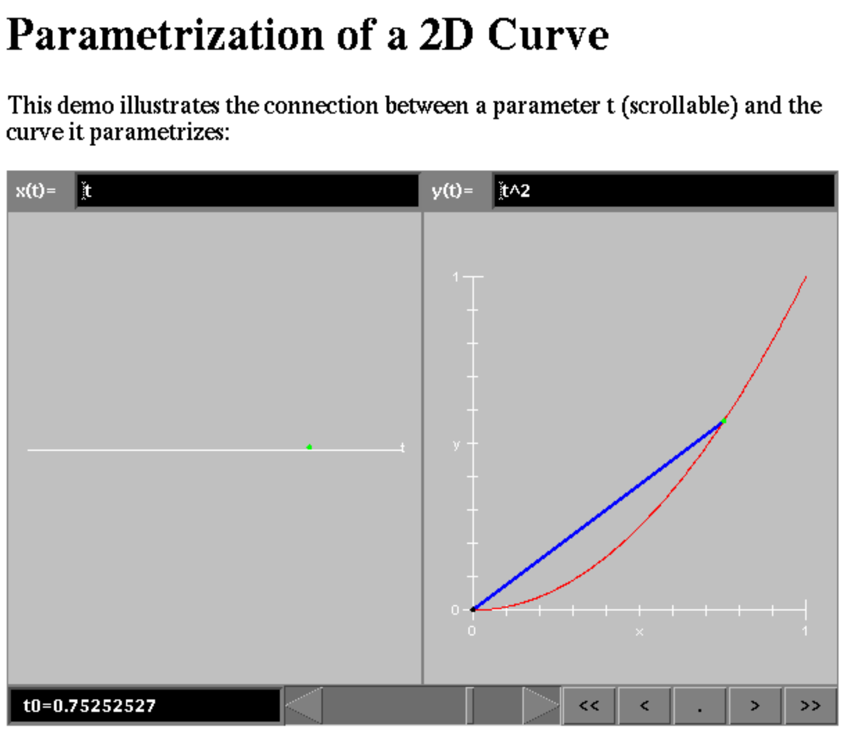


FIGURE 2. Parametrization of a plane curve

The curve C consists thus of all the points P_t "pointed to" by arrows $\overrightarrow{OP_t} = \mathbf{r}(t)$. Why do we insist on the *regularity condition* (2.5)? This is the subject of the next section.

EXAMPLE 2.11.
(*important general* example): Is it always possible to represent a curve $C \subset \mathbf{E}^2$, that is the *graph of a function* $f : I \to \mathbf{R}$, by a (vector function) parametrization? Yes! Here is the *recipe*:

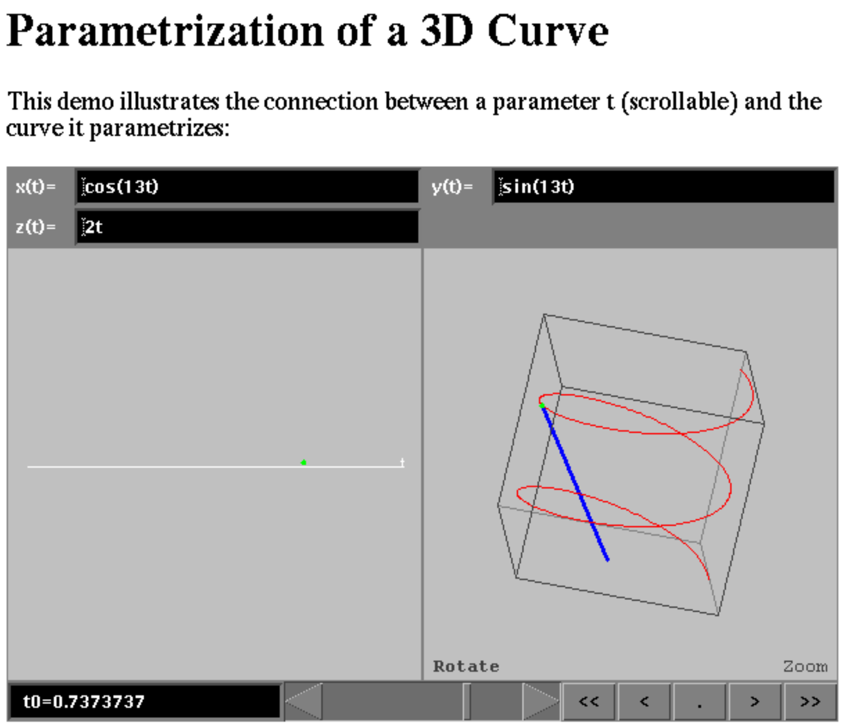


FIGURE 3. Parametrization of a space curve

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, $P_t : [t, f(t)]$ runs through all the points on the curve C. For instance, the graph of the sine-function has parametrization $\mathbf{r}(t) = [t, \sin t]$.

2.2. Tangents. Let $\mathbf{r} : [a, b] \to \mathbf{R}^i,\ i = 2$ or 3, denote a regular parametrization of a curve C with P_t corresponding to the parameter value t, i.e., $\mathbf{r}(t) = \overrightarrow{OP_t}$. Fix a point $P_{t_0} \in C$. Any line through P_{t_0} and another point P_t is called a *secant line* through P_{t_0}. What happens to those secant lines when t approaches t_0, and thus P_t approaches P_{t_0}? We shall show that, for a curve with a *regular* parametrization, there is a limit position: the *tangent line* through P_{t_0}[1].

The secant line through P_{t_0} and P_t has $\overrightarrow{P_{t_0}P_t}$ as a parallel vector. The following definition tells us what we mean by a limit position:

[1]Illustrations: next pages

DEFINITION 2.12. The positive, resp. negative *unit semi-tangent vectors* to the curve C at P_{t_0} are given by

$$\mathbf{t}_+(t_0) = \lim_{t \to t_0+} \frac{\overrightarrow{P_{t_0}P_t}}{|\overrightarrow{P_{t_0}P_t}|} \text{ resp. } \mathbf{t}_-(t_0) = \lim_{t \to t_0-} \frac{\overrightarrow{P_{t_0}P_t}}{|\overrightarrow{P_{t_0}P_t}|}.$$

If $\mathbf{t}_-(t_0) = -\mathbf{t}_+(t_0)$, then C has a tangent line with $\mathbf{t}(t_0) = \mathbf{t}_+(t_0)$ as a parallel vector.

REMARK 2.13.

(1) The limits of the vector functions above may be taken coordinatewise.
(2) In general, the limits above need not exist. But the following result shows, that we do not need to worry for curves with *regular* parametrizations:

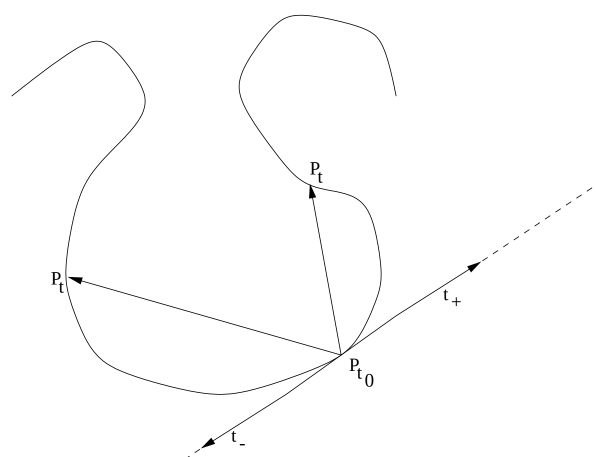


FIGURE 4. Secants and tangents through P_{t_0}

PROPOSITION 2.14. *Let* $\mathbf{r} : [a, b] \to \mathbf{R}^i$, $i = 2$ *or* 3, *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* $\mathbf{r}'(t_0)$. *In particular,*

$$\mathbf{r}'(t_0) = |\mathbf{r}'(t_0)|\mathbf{t}(t_0) \textit{ and } \mathbf{t}(t_0) = \frac{\mathbf{r}'(t_0)}{|\mathbf{r}'(t_0)|}.$$

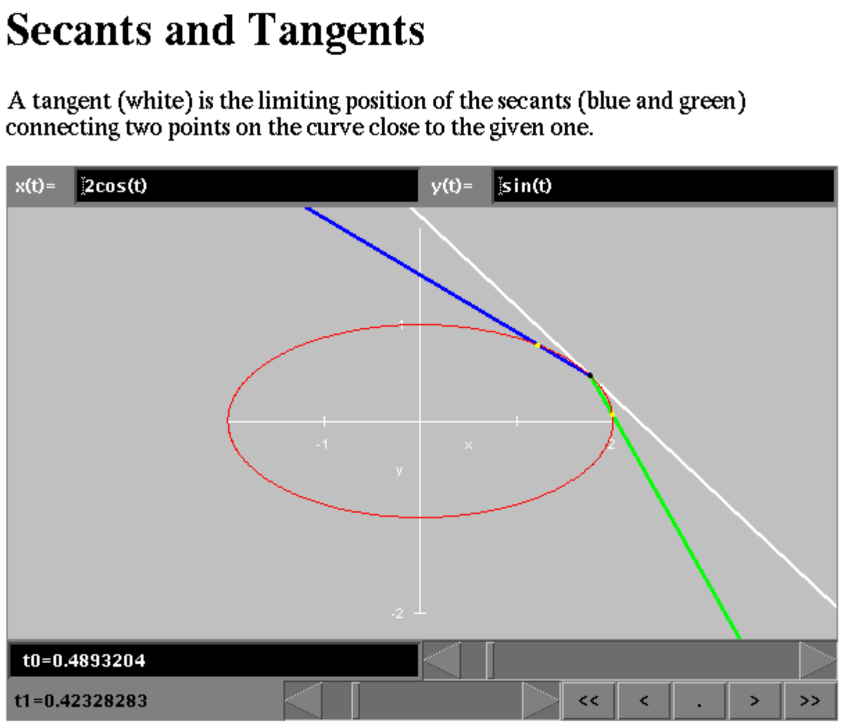


FIGURE 5. Secants and tangent

PROOF:
The scheme

$$\begin{array}{ccccc}
\frac{\mathbf{r}(t)-\mathbf{r}(t_0)}{t-t_0} & = & \frac{|\mathbf{r}(t)-\mathbf{r}(t_0)|}{t-t_0} & \cdot & \frac{\mathbf{r}(t)-\mathbf{r}(t_0)}{|\mathbf{r}(t)-\mathbf{r}(t_0)|} \\
{\scriptstyle t\to t_0+}\big\downarrow & & {\scriptstyle t\to t_0+}\big\downarrow & & {\scriptstyle t\to t_0+}\big\downarrow \\
\mathbf{r}'(t_0) & = & |\mathbf{r}'(t_0)| & \cdot & \mathbf{t}_+(t_0)
\end{array}$$

leads from an obvious equation between difference quotients to the equation asked for when t tends to t_0 for $t > t_0$. The latter assumption is used to ensure that $t - t_0 = |t - t_0|$. For $t < t_0$, the limiting equation becomes:

$$\mathbf{r}'(t_0) = -|\mathbf{r}'(t_0)| \cdot \mathbf{t}_-(t_0).$$

If $\mathbf{r}'(t_0) \neq \mathbf{0}$ – this is the case for a regular parametrization – we may calculate the unit tangent vector by division with $|\mathbf{r}'(t_0)|$.

□

A different interpretation of the tangent to a curve with a regular parametrization is given by zooming: If you zoom in closer and closer to the curve around a given point, the picture will become more and more linear: Close to that point, the curve and its tangent are almost the same; see the illustration on the opposite page.

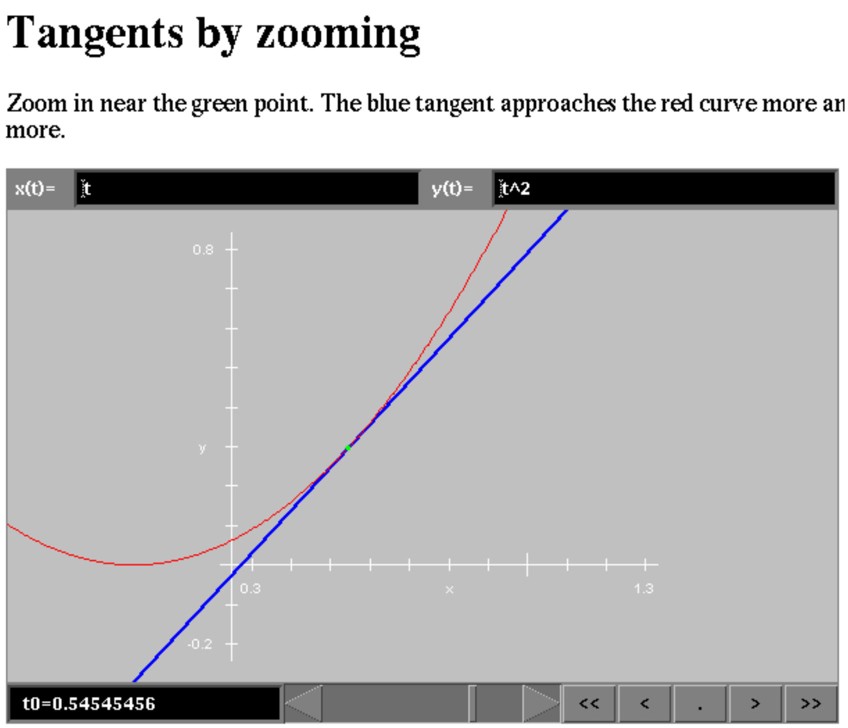


FIGURE 6. Tangents by zooming

For a regular parametrization of a curve C, the *velocity vector* function $\mathbf{v} = \mathbf{r}'$, contains two sorts of information: at every $t_0 \in (a, b)$, its *direction* determines the *tangent line* to the curve C at P_{t_0} – which is *independent* of the particular parametrization; its *magnitude* $|\mathbf{r}'(t_0)| = \lim_{t \to t_0} \frac{|\mathbf{r}(t) - \mathbf{r}(t_0)|}{|t - t_0|}$ corresponds to the *speed* of that *particular* parametrization:

DEFINITION 2.15. The vector $\mathbf{r}'(t_0)$ is called the *velocity vector* to C at P_{t_0}, the vector $\mathbf{t}(t_0) = \frac{\mathbf{r}'(t_0)}{|\mathbf{r}'(t_0)|}$ is called the *unit tangent vector* to C at P_{t_0}, the number $v(t_0) = |\mathbf{r}'(t_0)|$ is called the *speed* of the parametrization $\mathbf{r}$ at P_{t_0}.
In particular, $\mathbf{r}'(t_0) = v(t_0)\mathbf{t}(t_0)$.

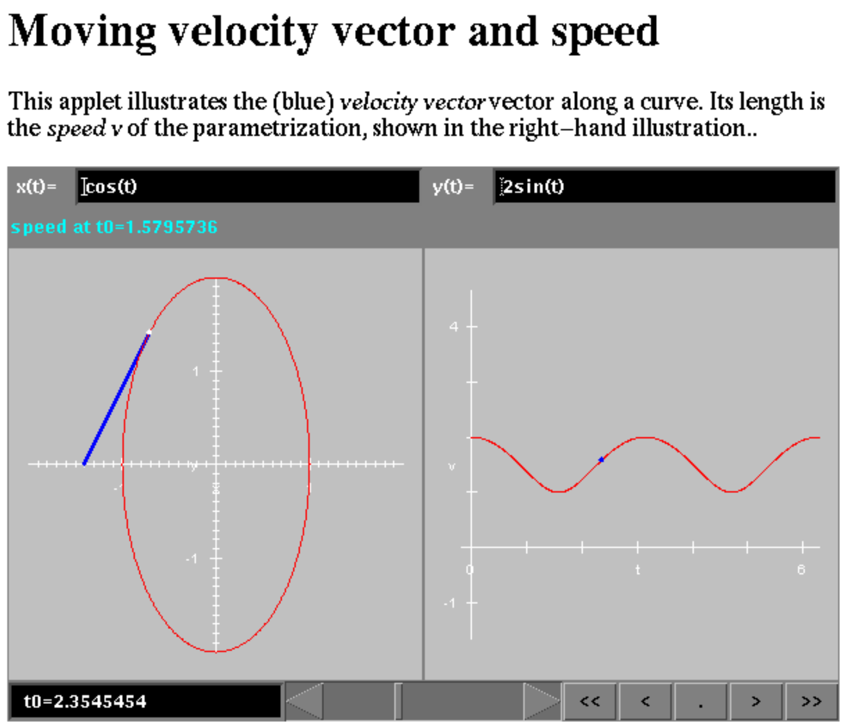


FIGURE 7. Moving velocity vector and speed

EXAMPLE 2.16.
The vector functions

$$\mathbf{r}_1 : [0, 2\pi] \longrightarrow \mathbf{R}^2, \quad \mathbf{r}_1(t) = [\cos t, \sin t]$$
$$\mathbf{r}_2 : [0, \sqrt{2\pi}] \longrightarrow \mathbf{R}^2, \quad \mathbf{r}_2(t) = [\cos(t^2), \sin(t^2)]$$

are both parametrizations of the unit circle around the origin in the plane, cf. Fig. 8. In fact $\mathbf{r}_2$ is a *reparametrization* of $\mathbf{r}_1$ since $\mathbf{r}_2(t) = \mathbf{r}_1(t^2)$.

At P_α with $\overrightarrow{OP_\alpha} = [\cos\alpha, \sin\alpha]$, the tangent vector is $\mathbf{t}_\alpha = [-\sin\alpha, \cos\alpha]$. The speed corresponding to $\mathbf{r}_1$ is 1 and in particular constant; the speed corresponding to $\mathbf{r}_2$ is $|\mathbf{r}_2'(t_0)| = 2t_0|[-\sin t_0^2, \cos t_0^2]| = 2t_0$, and thus not constant. At P_α – corresponding to $t_0 = \sqrt{\alpha}$, the speed is $2\sqrt{\alpha}$. Remark, that the parametrization $\mathbf{r}_2$ is *not* regular at $t_0 = 0$.

REMARK 2.17.
In general, let $\mathbf{r}_1 : [a, b] \longrightarrow \mathbf{R}^i, \mathbf{r}_2 : [c, d] \longrightarrow \mathbf{R}^i$ denote parametrizations for a curve. They have the same image (and run through it in the same direction) if there exists a strictly increasing differentiable function $f : [c, d] \longrightarrow [a, b]$ (with $f(c) = a$ and $f(d) = b$; a reparametrization) such that $\mathbf{r}_2(t) = \mathbf{r}_1(f(t))$.

The example above shows, that smooth curves (with tangent lines at every point) may have parametrizations that are not regular; moreover, there are important curves

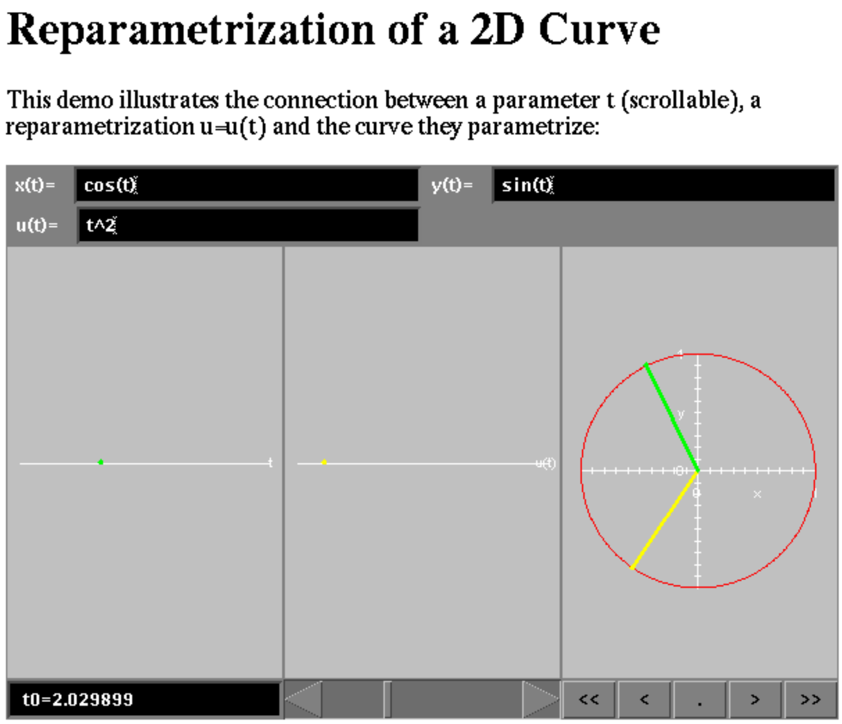


FIGURE 8. The illustration Reparametrization of a plane curve from the geometric laboratory GEOLAB allows you to generate many parametrizations of the same plane curve.

that *cannot* have a parametrization that is regular at *every* point. In order to get results on tangents for such curves, you have to calculate semi-tangent vectors at the points where the parametrization fails to be regular. This is illustrated in the following example:

EXAMPLE 2.18.
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. 9. We calculate: $\mathbf{r}'(t) = [2t, 3t^2]$, so the curve has a tangent at every point P_t for $t \neq 0$ according to Prop. 2.14. At P_0, $\mathbf{r}'(0) = \mathbf{0}$; nevertheless, one may calculate the semi-tangent vectors at this point:

$$\mathbf{t}_+(0) = \lim_{t \to 0+} \frac{\mathbf{r}(t) - \mathbf{r}(0)}{|\mathbf{r}(t) - \mathbf{r}(0)|} = \lim_{t \to 0+} \frac{t^2[1,t]}{t^2\sqrt{1+t^2}} = [1,0].$$

The calculation of $\mathbf{t}_-(0)$ yields the *same* result: $\mathbf{t}_+(0) = \mathbf{t}_-(0) = [1,0]$. Hence, the curve C *does not possess* a tangent line at (the singular point) $P_0 : [1,1]$, but two identical semi-tangents – in that situation the curve is said to have a “cusp” at the singular point. Hence, *no* parametrization of C can be regular. This corresponds to the following intuition: A particle moving along C will have to “jam on the brakes” reaching P_0 before backing and accelerating again.

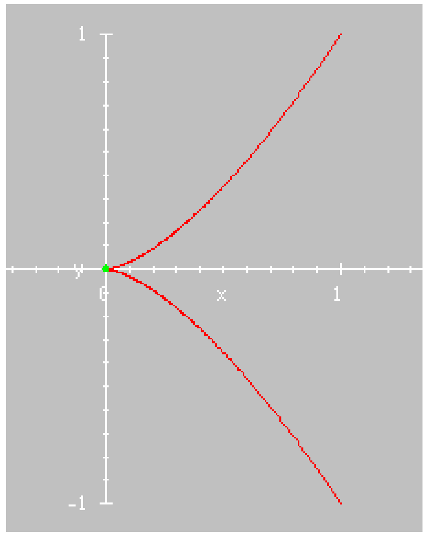


FIGURE 9. A curve with a singular point (a ″cusp″)

EXAMPLE 2.19.
A circular disk of radius 1 in the $xy-$plane rolls without slipping along the x-axis. The figure described by a point of the circumference of the disk is called a *cycloid* – imagine for example a point on the wheel of a bike.

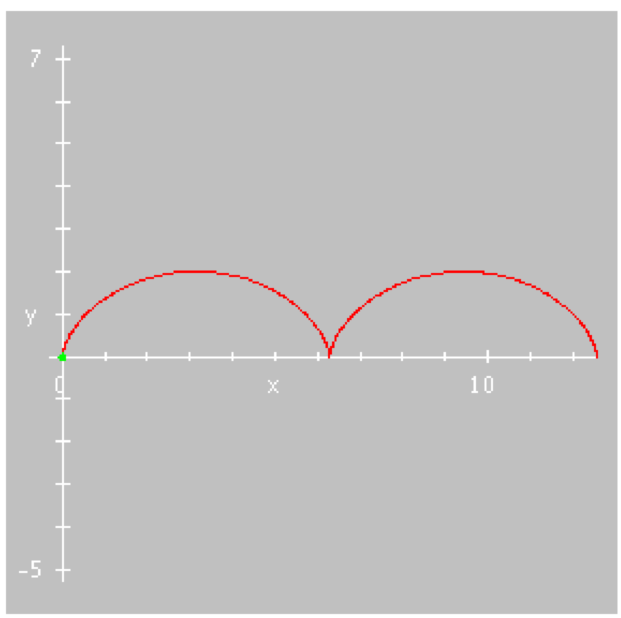


FIGURE 10. A cycloid

A parametrization for the cyloid is given as

$$\mathbf{r} : \mathbf{R} \to \mathbf{R}^2,\ \mathbf{r}(t) = [t - \sin t, 1 - \cos t].$$

It has singularities for $t = 2k\pi, k \in \mathbf{Z}$.

2.3. Arc Length.

2.3.1. *Definition and Calculation.* Let again C be a curve with a regular parametrization $[a, b] \to \mathbf{R}^i, i = 2$ or $i = 3$. What is the length l of the piece of curve between the starting point P_a and a point $P_t,\ t \in [a, b]$? What is the length of the entire curve from P_a to P_b?

Let the length of the piece of curve from P_a to P_t be denoted by $s(t)$. Of course, $s(a) = 0$, and s is an increasing (ordinary) function on $[a, b]$. Moreover, we found the *speed* of the parametrization as the function $v : [a, b] \to \mathbf{R}$, $v(t) = |\mathbf{r}'(t)|$. The speed is the differential increment of the length:

$$v(t) = \lim_{h \to 0} \frac{s(t+h) - s(t)}{h} = s'(t), \text{ and hence:}$$

DEFINITION 2.20. The *arc length* function $s : [a, b] \to \mathbf{R}$ corresponding to the curve C parametrized by the vector function $\mathbf{r}$ above is defined as

$$s(t_0) = \int_a^{t_0} v(t)dt = \int_a^{t_0} |\mathbf{r}'(t)|dt.$$

More explicitly, if $\mathbf{r}(t) = [x(t), y(t), z(t)]$, then

$$s(t_0) = \int_a^{t_0} \sqrt{(x')^2(t) + (y')^2(t) + (z')^2(t)}dt.$$

The length of the entire curve C is

$$l = s(b) = \int_a^b |\mathbf{r}'(t)|dt = \int_a^b \sqrt{(x')^2(t) + (y')^2(t) + (z')^2(t)}dt.$$

The Fundamental Theorem of Calculus allows to calculate the derivative of the arc length function $s(t)$: In fact, $s' = v$, the speed, as it should! Furthermore, one can show – using integration by substitution – that the definition of the arc length function above is *independent* of the chosen parametrization $\mathbf{r}$.

EXAMPLE 2.21.

(1) For the helix from Ex. 2.9 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. 2.11 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!

2.3.2. *Parametrization by Arc Length.* Let C be a smooth regular curve in the plane or in space. We want to single out a particular parametrization for c:

DEFINITION 2.22. The *parametrization by arc length* $\mathbf{r}_{al} : [0, l] \to \mathbf{R}^i$, $i = 2$ or $i = 3$, is characterized by the property:
$|\mathbf{r}'_{al}(s)| = 1$ for all $s \in (0, l)$.

In other words:

- The speed of this particular parametrization is 1 and, in particular, constant;
- the derivative of $\mathbf{r}_{al}$ is the *unit* tangent vector $\mathbf{t}(s)$ at every element $s \in (0, l)$: $\mathbf{t}(s) = \mathbf{r}'_{al}(s)$.

The arc length parametrization has the following property which explains its name:

PROPOSITION 2.23.

$$\int_0^{s_0} |\mathbf{r}'_{al}(s)| ds = s_0 \textit{ for every } s_0 \in [0, l],$$

i.e., the parameter s is equal to the length of the piece of curve (arc) from its start point $P(0)$ *with* $\overrightarrow{OP(0)} = \mathbf{r}_{al}(0)$ *to the point* $P(s)$ *with* $\overrightarrow{OP(s)} = \mathbf{r}_{al}(s)$.

EXAMPLE 2.24.

The parametrization $\mathbf{r}(t) = [a\cos t, a\sin t, bt]$ for the helix from Ex. 2.9 yields (cf. Ex. 2.21): $v(t) = |\mathbf{r}'(t)| = \sqrt{a^2 + b^2}$; 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 $t(s) = \frac{s}{\sqrt{a^2+b^2}}$. 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 2.25.

(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) In most cases, it is *not* possible to write down an *explicit* formula for the arc length function corresponding to a curve. Given a regular parametrization $\mathbf{r}(t)$ as above, one may argue for the existence of an inverse function $t(s)$ of $s(t)$ and substitute: $\mathbf{r}_{al}(s) = \mathbf{r}(t(s))$. But the inverse function $t(s)$ can in most cases not be calculated explicitly. Just try for the graph of a differentiable function (cf. Ex. 2.11).

(3) 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. In this way, 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.

The connection between an arbitrary regular parametrization $\mathbf{r}$ of a curve C and its arc length parametrization $\mathbf{r}_{al}$ is given by the formula

$$\mathbf{r}(t) = \mathbf{r}_{al}(s(t)). \tag{2.6}$$

(Read: At time t, the vector $\mathbf{r}(t)$ points to the same point as $\mathbf{r}_{al}(s(t))$ (at distance $s(t)$ from the beginning point along the curve). Differentiating this equation using the chain rule 2.4.5, we obtain the following expression for the velocity vector $\mathbf{v}(t)$:

$$\mathbf{v}(t) = \mathbf{r}'(t) = s'(t)\mathbf{r}'_{al}(s(t)) = v(t)\mathbf{t}(t), \tag{2.7}$$

i.e., the velocity vector has length $v(t)$ – the speed – and is parallel to the unit tangent vector $\mathbf{t}(t)$.

3. Curvature

What is the curvature of a curve? 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.

In Sect. 3.2.1, we shall see, that the curvature of a curve gives information about the *normal* component of the *acceleration vector* of a particle moving along the curve. This is one of the reasons for its importance in mechanical applications.

3.1. Definitions. Let C be a regular smooth curve in plane or in space with arc length parametrization $\mathbf{r}_{al} : [0, l] \to \mathbf{R}^i$. By Def. 2.22, $\mathbf{t}(s) = \mathbf{r'}_{al}(s)$ is a unit tangent vector to the curve at the point P_s. The vector function $\mathbf{t} : [0, l] \to \mathbf{R}^i$ 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 $\mathbf{t}' = \mathbf{r''}_{al}$ along the curve C. An application of Prop. 2.5 yields:

COROLLARY 2.26. (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.

PROOF. (1) $|\mathbf{t}(s)| = 1 \Leftrightarrow \mathbf{t}(s) \cdot \mathbf{t}(s) = 1$. Apply Prop. 2.5.
(2) $\mathbf{t}'(s)$ and $\hat{\mathbf{t}}(s)$ are both perpendicular to $\mathbf{t}(s)$ and hence parallel to each other. □

DEFINITION 2.27. (1) The vector $\mathbf{t}'(s)$ is called the *curvature vector*[a] at the point P_s on C with $\overrightarrow{OP_s} = \mathbf{r}_{al}(s)$.

(2) 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 with $\mathbf{t}'(s) \neq \mathbf{0}$ let $\mathbf{n}(s) = \frac{\mathbf{t}'(s)}{|\mathbf{t}'(s)|}$.

(3) For a plane curve C, the pair $[\mathbf{t}(s), \mathbf{n}(s)]$ is called the *moving frame* along C[b].

[a]See illustration below

[b]See illustration on opposite page

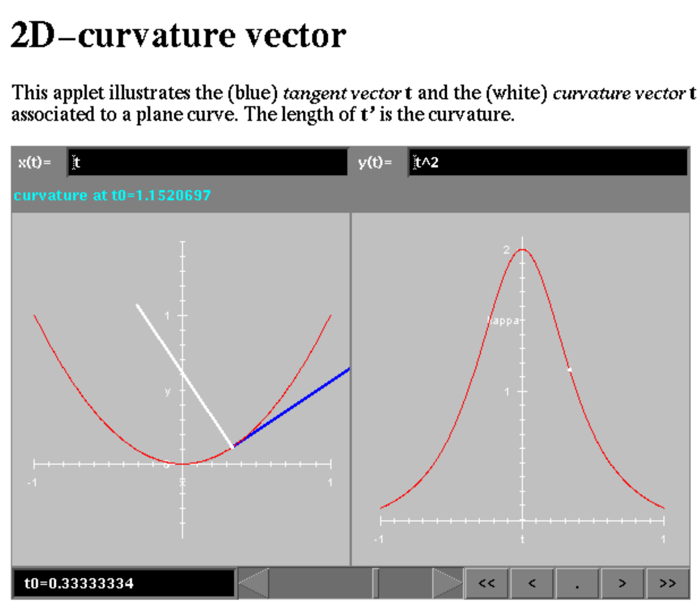


FIGURE 11. Curvature vector along a space curve

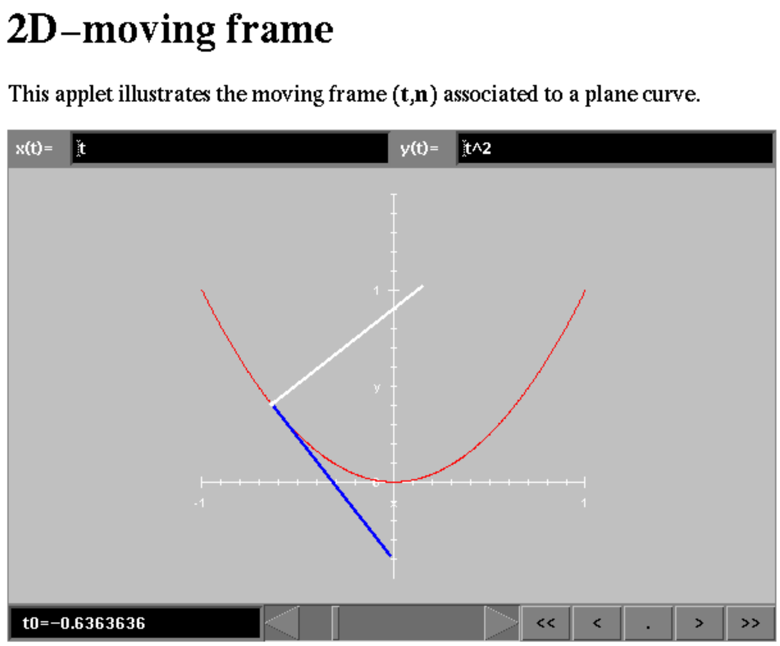


FIGURE 12. Moving frame for a plane curve

DEFINITION 2.28. The *curvature* $\kappa(P_s)$ of the curve C at the point P_s is defined as follows[a]:

(1) For a *plane* curve C let

$$\mathbf{t}'(s) = \kappa(s)\mathbf{n}(s), \tag{2.8}$$

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

(2) For a *space* curve C let

$$\kappa(s) = |\mathbf{t}'(s)| \geq 0. \tag{2.9}$$

[a]Illustrations: following pages

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

(1) Near a point P on the curve C, that curve may be approximated by an *approximating circle* through nearby points P_0 and P_1 on the curve C – if the three points are not contained in a line. If P_0 and P_1 both get closer and closer (converge) to P, this approximating circle converges to the *osculating circle* of C at P, cf. Sect. 3.4 and Fig. 15 and 16. The curvature $\kappa(s)$ of the curve C at P is in inverse proportion to the radius $r(s)$ of this osculating circle: $\kappa(s) = \frac{1}{r(s)}$.

2D–Curvature

This applet illustrates the curvature of a plane curve:

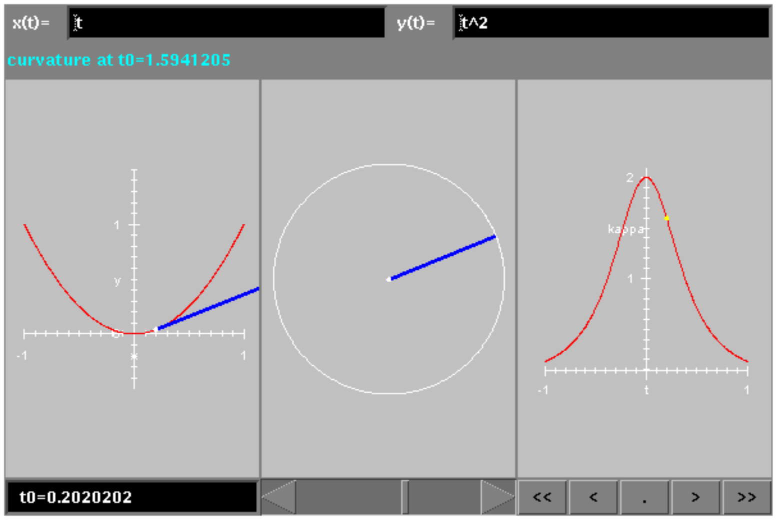


FIGURE 13. Gauss map and curvature along a plane curve

3D–Curvature

This applet illustrates the curvature of a space curve:

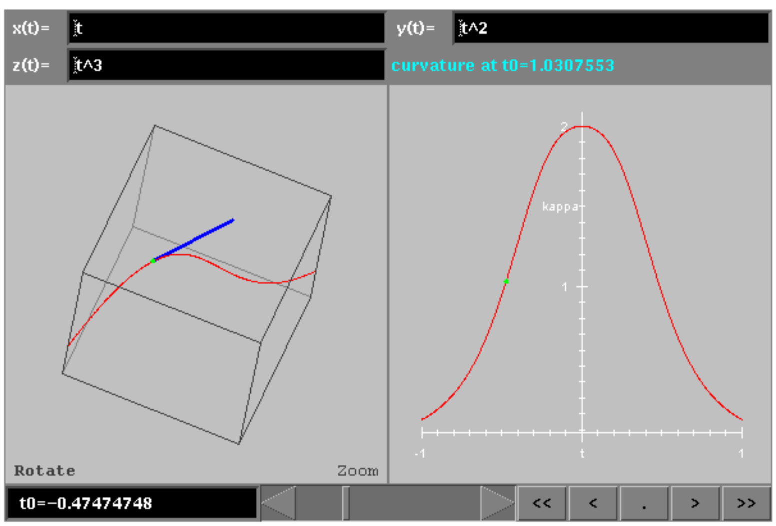


FIGURE 14. Curvature along a space curve

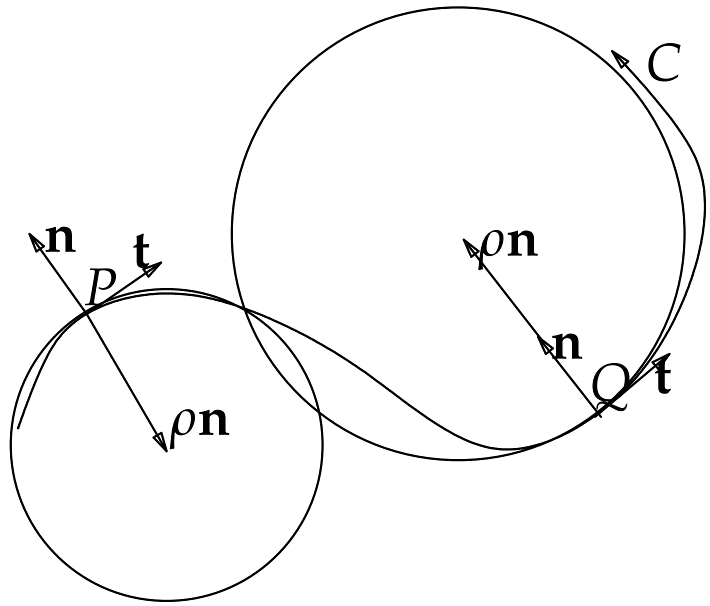


FIGURE 15. Tangent and principal normal vectors and osculating circles at points P and Q

(2) 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 speed is constant. The faster the direction of the tangent vector changes, the more curved is the curve.

(3) In the plane, we can parameterize every unit vector as $[\cos\theta, \sin\theta]$, and hence the unit tangent vector field $\mathbf{t}(s) = [\cos\theta(s), \sin\theta(s)]$ with $\theta(s)$ the angle between $\mathbf{t}(s)$ and the horizontal vector $\mathbf{i}$. The map $s \mapsto \theta(s)$ with values on the unit circle in the plane is also called the *Gauss map* associated to the curve, cf. Fig. 13. A calculation of $\mathbf{t}'(s)$ using the chain rule (Prop. 2.4(5)) yields:

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

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 explanantion for the *sign* of the curvature of a plane curve, to wit:

COROLLARY 2.29. *Near the point P_s, a plane curve C is curved*
- counter-clockwise *if and only if $\kappa(s) > 0$,*
- *and* clockwise *if and only if $\kappa(s) < 0$.*

In Fig. 15, the curvature is negative at P and positive at Q.

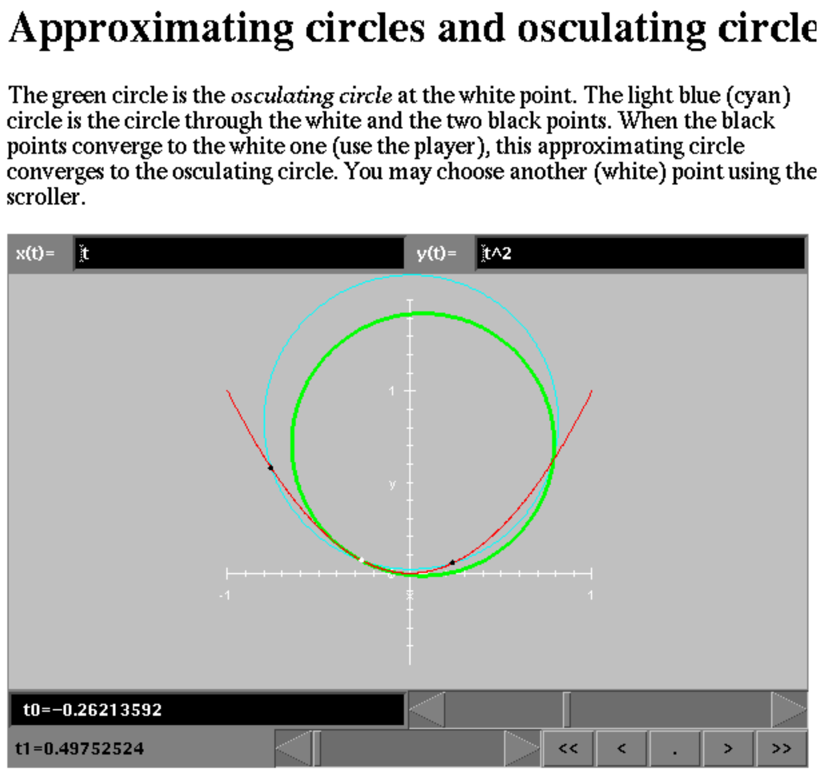


FIGURE 16. Approximating circles and osculating circle

(4) 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 (2.8), $\kappa(s) = \frac{1}{R}$ 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. 15 and 16 and Sect. 3.4.

3.2. Calculation of the Curvature.

3.2.1. *Components of the acceleration vector.* The definition of curvature above uses the arc length parametrization of a given curve. But in general, you have only a *regular* parametrization $\mathbf{r} : [a, b] \to \mathbf{R}^i$ at hand. In this case, the curvature at a given point P is "hidden" in the *acceleration vector* at that point:

With respect to the chosen parametrization $\mathbf{r}(t)$, the acceleration vector $\mathbf{a}(t)$ is the derivative of the velocity vector $\mathbf{v}(t)$:

$$\mathbf{a}(t) = \mathbf{v}'(t) = \mathbf{r}''(t).$$

We use the expression

$$\mathbf{v}(t) = \mathbf{r}'(t) = v(t)\mathbf{t}(t)$$

from (2.7) to calculate the *acceleration* vector $\mathbf{a}(t)$ using both the product and the chain rule from Prop. 2.4.2 and .5:

$$\mathbf{a}(t) = \mathbf{v}'(t) = v'(t)\mathbf{t}(t) + v(t)\mathbf{t}'(t) = v'(t)\mathbf{t}(t) + v^2(t)\kappa(t)\mathbf{n}(t) \tag{2.10}$$

with $\mathbf{n}(t)$ as in (2.28). The second factor $v(t)$ in (2.10) is explained by the chain rule (cf. Prop. 2.4.5), since $\mathbf{r}(t) = \mathbf{r}_{al}(s(t))$, and $(\mathbf{t} \circ s)'(t) = s'(t)\mathbf{t}'(s(t)) = v(t)\kappa(t)\mathbf{n}(t)$.

Before using (2.10) to calculate the curvature of a given curve, let us look at the following attractive *interpretation in mechanics*: Equation (2.10) yields a decomposition of the acceleration vector $\mathbf{a}(t)$ into a *tangential* component $\mathbf{a_t}(t)$ and a *normal* component $\mathbf{a_n}(t)$:[2]

$$\mathbf{a}(t) = \mathbf{a_t}(t) + \mathbf{a_n}(t) = v'(t)\mathbf{t}(t) + v^2(t)\kappa(t)\mathbf{n}(t). \tag{2.11}$$

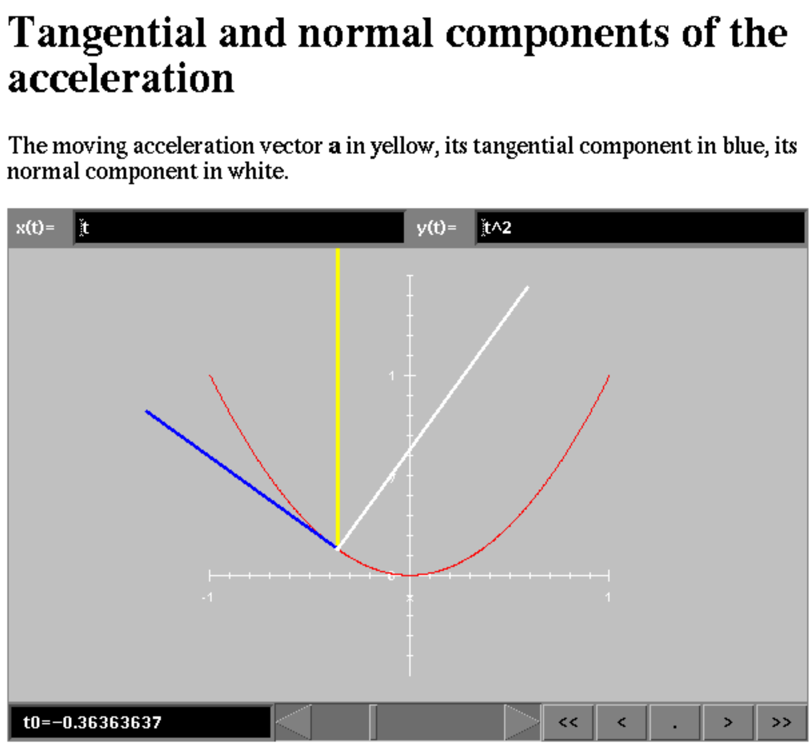


FIGURE 17. Tangential and normal components of the acceleration vector

[2]illustration: next page

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. Every engineer planning roads or railways has to know this very explicitly!

Remark that (2.11) can be interpreted as to yield the *projections* of the acceleration vector $\mathbf{a}(t)$ on the unit tangent vector $\mathbf{t}(t)$, resp. on the principal normal vector $\mathbf{n}(t)$. Using the formulae from (1.26) and (1.27) we can calculate these projections and their lengths using just the velocity vector and the acceleration vector:

$$\mathbf{a_t}(t) = \frac{\mathbf{r}''(t) \cdot \mathbf{r}'(t)}{|\mathbf{r}'(t)|^2}\mathbf{r}'(t),\ |\mathbf{a_t}(t)| = \frac{\mathbf{r}''(t) \cdot \mathbf{r}'(t)}{|\mathbf{r}'(t)|},$$

$$\text{(plane) } \mathbf{a_n}(t) = \frac{[\mathbf{r}'(t), \mathbf{r}''(t)]}{|\mathbf{r}'(t)|^2}\hat{\mathbf{r}}'(t),\ |\mathbf{a_n}(t)| = \frac{|[\mathbf{r}'(t), \mathbf{r}''(t)]|}{|\mathbf{r}'(t)|}$$

$$\text{(space) } \mathbf{a_n}(t) = \frac{1}{|\mathbf{r}'(t)|^2}\mathbf{r}'(t) \times (\mathbf{r}''(t) \times \mathbf{r}'(t)),\ |\mathbf{a_n}(t)| = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|}.$$

3.3. Curvature formulas.

PROPOSITION 2.30. (1) *Let C be a* plane *curve with parametrization* $\mathbf{r} : [a,b] \to \mathbf{R}^2$. *Its curvature* $\kappa(P)$ *at a point P with* $\overrightarrow{OP} = \mathbf{r}(t)$ *is given by*

$$\kappa(P) = \frac{[\mathbf{r}'(t), \mathbf{r}''(t)]}{|\mathbf{r}'(t)|^3}. \tag{2.12}$$

The denominator is the plane product ((1.8) in Sect. 1.2.2) *of the vectors* $\mathbf{r}(t)$ *and* $\mathbf{r}'(t)$.

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

$$\kappa(P) = \frac{\begin{vmatrix} x'(t) & x''(t) \\ y'(t) & y''(t) \end{vmatrix}}{(\sqrt{x'(t)^2 + y'(t)^2})^3}. \tag{2.13}$$

(2) *Let C be a* space *curve with parametrization* $\mathbf{r} : [a,b] \to \mathbf{R}^3$. *Its curvature* $\kappa(P)$ *at a point P with* $\overrightarrow{OP} = \mathbf{r}(t)$ *is given by*

$$\kappa(P) = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|^3}. \tag{2.14}$$

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

$$\kappa(P) = \frac{|[x'(t), y'(t), z'(t)] \times [x''(t), y''(t), z''(t)]|}{(\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2})^3}. \tag{2.15}$$

PROOF:

(1) To extract $\kappa(t)$ from (2.10), we calculate the dot product of both sides of that equation with the vector $\widehat{\mathbf{r}'}(t)$. In the equations below We exploit that thus vector equals $v(t)\hat{\mathbf{t}}(t)$ according to (2.7) and to $v(t)\mathbf{n}(t)$ according to Definition 2.27(2a):
$\mathbf{r}''(t) \cdot \widehat{\mathbf{r}'}(t) = (v'(t)\mathbf{t}(t) + v^2(t)\kappa(t)\mathbf{n}(t)) \cdot v(t)\mathbf{n}(t) = v^3(t)\kappa(t)(\mathbf{n}(t) \cdot \mathbf{n}(t)) = v^3(t)\kappa(t)$.
Using the plane product from Section 1.2.2 and its properties from Section 1.2.6, we obtain:
$[r'(t), r''(t)] = \mathbf{r}''(t) \cdot \widehat{\mathbf{r}'}(t)$ and hence: $\kappa(t) = \frac{1}{v^3(t)}[r'(t), r''(t)] = \frac{[\mathbf{r}'(t), \mathbf{r}''(t)]}{|\mathbf{r}'(t)|^3}$.

(2) Using properties of the cross product and (2.10), 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 2.31.

(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 have:

$$\kappa(P_t) = \frac{ab}{(a^2(\sin t)^2 + b^2(\cos t)^2)^{\frac{3}{2}}}. \qquad (2.16)$$

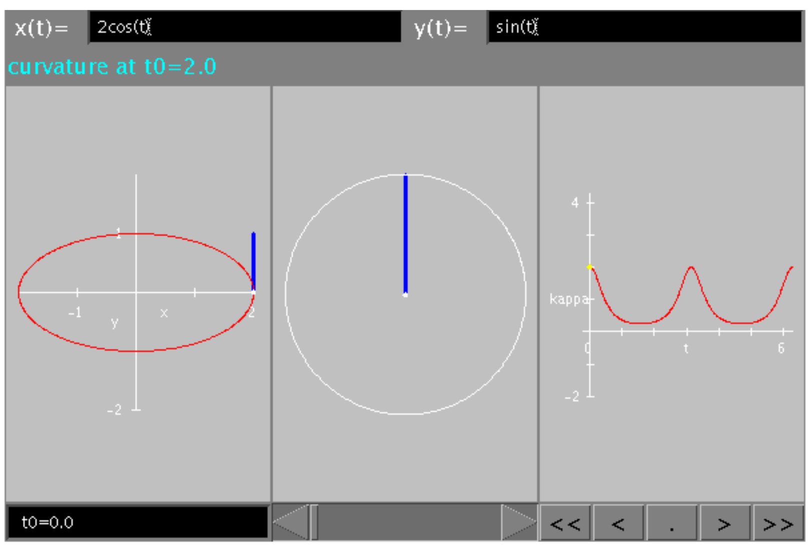


FIGURE 18. Ellipse and curvature function

Fig. 18 shows (for $a = 2$ and $b = 1$) how the curvature oscillates periodically along the ellipse – with extremal values in the points on the axes.

(2) The curve C_f given as the *graph* of a function $f : [a, b] \to \mathbf{R}$ (Ex. 2.11) 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

$$\kappa(P_t) = \frac{f''(t)}{\sqrt{1+(f'(t))^2}^3}. \tag{2.17}$$

In particular, C_f is curved counter-clockwise at P_t if $f''(t) > 0$ and clockwise if $f''(t) < 0$ (Cor. 2.29).

(3) Let $\mathbf{r}$ denote the parametrization for a helix from Ex. 2.9 given as

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

Its derivatives (velocity and acceleration vectors) are calculated as

$$\begin{aligned} \mathbf{r}'(t) &= [-a\sin t, a\cos t, b]; \\ \mathbf{r}''(t) &= [-a\cos t, -a\sin t, 0]. \end{aligned} \tag{2.18}$$

Hence,

$$\begin{aligned} |\mathbf{r}'(t)| &= \sqrt{a^2+b^2}; \\ \mathbf{r}'(t)\times\mathbf{r}''(t) &= [ab\sin t, -ab\cos t, a^2]; \\ |\mathbf{r}'(t)\times\mathbf{r}''(t)| &= a\sqrt{a^2+b^2}. \end{aligned} \tag{2.19}$$

The curvature of the helix at P_0 with $\overrightarrow{OP_0} = \mathbf{r}(t_0)$ is calculated as

$$\kappa(P_0) = \frac{a\sqrt{a^2+b^2}}{\sqrt{a^2+b^2}^3} = \frac{a}{a^2+b^2}. \tag{2.20}$$

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.

3.4. Osculating Circles and the Evolute. Let C denote a curve and P a point on it. The *tangent line* l_P to C at P is the best approximating line through P. In particular, the 1.degree Taylor polynomial to a parametrization $\mathbf{r}$ of C, cf. (2.3), yields for us a parametrization $\mathbf{r}_{t_0}^{(1)}$ of l_P with coinciding values of the vector functions and their first derivatives at t_0 corresponding to P:

$$\mathbf{r}(t_0) = \mathbf{r}_{t_0}^{(1)}(0);\ \mathbf{r}'(t_0) = \mathbf{r}'^{(1)}_{t_0}(0).$$

3.4.1. *Osculating Circles.* A better approximation of the curve C than by lines can be obtained by *osculating circles*[3]. We ask at every point $P \in C$ for *the* circle through P with the *same tangent vector* and *same curvature vector*, cf. Def. 2.28. In particular, their curvatures have to agree (up to a sign). This approximating circle is called the *osculating circle* of the curve C at P.[4]

[3] cf. Motivation 1. after Def. 2.28 and Fig. 19

[4] *Osculum* is the Latin word for *kiss*.

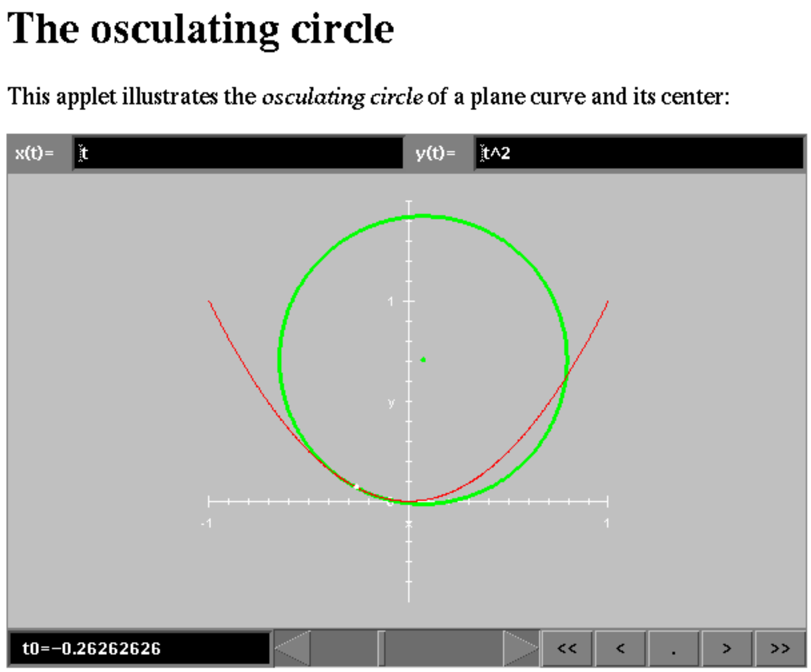


FIGURE 19. The osculating circle

To *find* the osculating circle, assume that the curve has parametrization by arc length $\mathbf{r}_{al}$ with $\mathbf{r}_{al}(s_0) = \overrightarrow{OP}$. The osculating circle has a parametrization $\mathbf{c}_{al}$ by arc length, too. It must have the form

$$\mathbf{c}_{al}(s) = \mathbf{c} + R\cos(\frac{s}{R})\mathbf{v}_1 + R\sin(\frac{s}{R})\mathbf{v}_2 \tag{2.21}$$

with unknown *orthogonal unit* vectors $\mathbf{v}_1, \mathbf{v}_2$, an unknown *center of curvature* C_P with $\overrightarrow{OC_P} = \mathbf{c}$. Assume $\mathbf{c}_{al}(0) = \overrightarrow{OP}$.

The radius R of the circle is in inverse proportion to its curvature, which in turn coincides with the curvature $\kappa_0 = \kappa(s_0)$ of the curve C at P, i.e., $R = \frac{1}{|\kappa_0|}$. We replace R with $\frac{1}{\kappa_0}$ in (2.21) – this makes sense even for negative curvature – and calculate:

$$\mathbf{c}_{al}(s) = \mathbf{c} + \frac{\cos(\kappa_0 s)}{\kappa_0}\mathbf{v}_1 + \frac{\sin(\kappa_0 s)}{\kappa_0}\mathbf{v}_2,$$

$$\mathbf{c}'_{al}(s) = -\sin(\kappa_0 s)\mathbf{v}_1 + \cos(\kappa_0 s)\mathbf{v}_2,$$

$$\mathbf{c}''_{al}(s) = -\kappa_0\cos(\kappa_0 s)\mathbf{v}_1 - \kappa_0\sin(\kappa_0 s)\mathbf{v}_2,$$

and argue as follows:

(1) Curve and circle pass through P:
$\overrightarrow{OP} = \mathbf{r}_{al}(s_0) = \mathbf{c}_{al}(0) = \mathbf{c} + \frac{1}{\kappa_0}\mathbf{v}_1.$

(2) Curve and circle have the same tangent vector at P :
$\mathbf{t}(s_0) = \mathbf{r}'_{al}(s_0) = \mathbf{c}'_{al}(0) = \mathbf{v}_2$.
(3) Curve and circle have the same curvature vector at P :
$\kappa(s_0)\mathbf{n}(s_0) = \mathbf{t}'(s_0) = \mathbf{r}''_{al}(s_0) = \mathbf{c}''_{al}(0) = -\kappa_0\mathbf{v}_1$, and hence:
$\mathbf{v}_1 = -\frac{1}{\kappa_0}\mathbf{r}''_{al}(s_0) = -\frac{1}{\kappa_0}\mathbf{t}'(s_0) = -\mathbf{n}(s_0)$.

From 1. and 3. above, we can conclude:

$$\overrightarrow{OC_P} = \mathbf{c} = \overrightarrow{OP} - \frac{1}{\kappa_0}\mathbf{v}_1 = \overrightarrow{OP} + \frac{1}{\kappa_0}\mathbf{n}(s_0).$$

Thus, we have proved:

PROPOSITION 2.32. *Let C be a regular curve through a point P with unit tangent vector $\mathbf{t}_0$, principal normal vector $\mathbf{n}_0$ and curvature κ_0 at P. The osculating circle through the point P on C has radius $R = \frac{1}{|\kappa_0|}$. Its center C_P is on the principal normal line through P and given by*

$$\overrightarrow{OC_P} = \overrightarrow{OP} + \frac{1}{\kappa_0}\mathbf{n}_0. \tag{2.22}$$

The osculating circle can be parameterised as follows:

$$\mathbf{c}_{al}(s) = (\overrightarrow{OP} + \frac{1}{\kappa_0}\mathbf{n}_0) - \frac{1}{\kappa_0}\cos(s\kappa_0)\mathbf{n}_0 + \frac{1}{\kappa_0}\sin(s\kappa_0)\mathbf{t}_0. \tag{2.23}$$

It is contained in the osculating plane (cf. Sec. 4.1) *through P spanned by the vectors $\mathbf{t}_0$ and $\mathbf{n}_0$.*

EXAMPLE 2.33.
Let E denote the ellipse from Ex. 2.31.1 with parametrization $\mathbf{r}(t) = [a\cos t, b\sin t]$. At P_0 with $\mathbf{r}(t_0) = \overrightarrow{OP_0}$, we obtain from (2.16) the radius of curvature

$$\rho_0 = \frac{1}{|\kappa(t_0)|} = \frac{(a^2(\sin t_0)^2 + b^2(\cos t_0)^2)^{\frac{3}{2}}}{ab}.$$

To calculate the center of curvature, we need the normal $\mathbf{n}_0$ at P_0. From Ex. 2.31.1, we extract:

$$\begin{aligned}
\mathbf{r}'(t_0) &= [-a\sin t_0, b\cos t_0];\\
\mathbf{t}(t_0) &= \frac{1}{(a^2(\sin t_0)^2+b^2(\cos t_0)^2)^{\frac{1}{2}}}[-a\sin t_0, b\cos t_0];\\
\mathbf{n}_0=\widehat{\mathbf{t}}(t_0) &= \frac{1}{(a^2(\sin t_0)^2+b^2(\cos t_0)^2)^{\frac{1}{2}}}[-b\cos t_0, -a\sin t_0];\\
\rho_0\mathbf{n}_0 &= \frac{a^2(\sin t_0)^2+b^2(\cos t_0)^2}{ab}[-b\cos t_0, -a\sin t_0]\\
&= [-a\cos t_0+\frac{a^2-b^2}{a}\cos^3 t_0, -b\sin t_0+\frac{b^2-a^2}{b}sin^3 t_0].
\end{aligned}$$

The last equation uses the trigonometric identity $\sin^2 t_0+\cos^2 t_0 = 1$. The osculating circle has its center at C_0 with $\overrightarrow{OC_0}=\overrightarrow{OP_0}+\rho_0\mathbf{n}_0=$
$=[a\cos t_0, b\sin t_0]+[-a\cos t_0+\frac{a^2-b^2}{a}\cos^3 t_0, -b\sin t_0+\frac{b^2-a^2}{b}sin^3 t_0]=$
$=[\frac{a^2-b^2}{a}\cos^3 t_0, \frac{b^2-a^2}{b}sin^3 t_0]$.

3.4.2. *The Evolute Curve.* From a given curve C, one may obtain a new curve E_C, the *evolute*[5] of C, by associating to every point $P\in C$ the corresponding *centre of curvature* C_P. For a *plane* curve C, it is easy to translate (2.22) into a parametrization of E_C:

COROLLARY 2.34. *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:*

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

It turns out, that the tangent line to the evolute at C_P coincides with the principal normal line to the curve C at P.

EXAMPLE 2.35.

(1) The evolute of a cycloid is a translated cycloid.
(2) The evolute of an ellipse is an *astroid*. Using Ex. 2.33 and/or Cor. 2.34, we obtain the parametrization $\mathbf{e}(t)=[\frac{a^2-b^2}{a}(\cos t)^3, \frac{b^2-a^2}{b}(\sin t)^3]$.

[5] illustration: next page

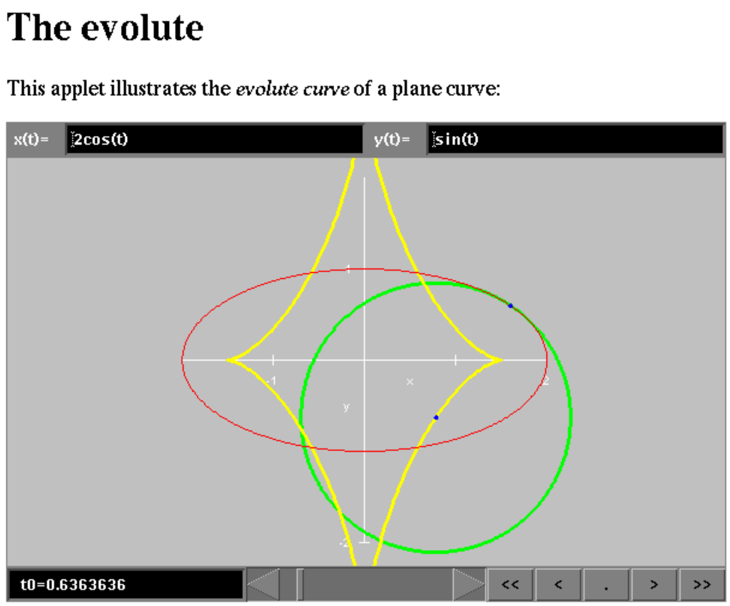


FIGURE 20. The evolute of an ellipse is an astroid (yellow)

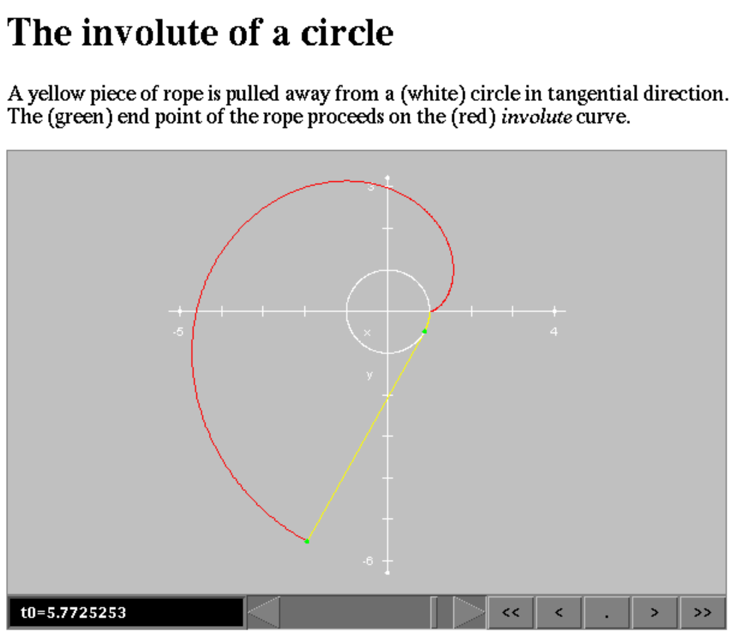


FIGURE 21. Reversing the evolution process, one arrives at still another curve associated to C, the *involute* of C. Above you find the involute of a circle.

3.5. The curvature function determines a plane curve. Let $\kappa : I = [0, l] \rightarrow \mathbf{R}$ denote a smooth function. Is there always a plane curve C (with arclength parametrization $\mathbf{r}_{al} : I \rightarrow \mathbf{R}^2$) realizing this function as the curvature function associated to C? How many such curves are there?

In fact, 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 there is a good intuitive explanation for the result above: 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 "little" circle 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"little" circle 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!

The correct solution uses the infinitesimal version of this idea. A curve with curvature function κ starting at the origin and in the direction of the X-axis may be given as

$$\mathbf{r}(s) = [\int_0^s \cos\theta(u)du, \int_0^s \sin\theta(u)du].$$

The function $\theta(u)$ measures the angle between the tangent and the X-axis and is determined form the curvature function by $\theta(u) = \int_0^u \kappa(t)dt$. Using the Fundamental Theorem of Calculus, it is not difficult to see that $\mathbf{r}$ has unit speed; differentiating $\mathbf{t}(s) = \mathbf{r}'(s)$ shows, that one obtains $\theta' = \kappa$ as the curvature.

This might seem 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. 22 is called a *clothoid* curve – it has $\kappa(s) = s$ as the associated curvature function, and you have probably experienced it as a driver!

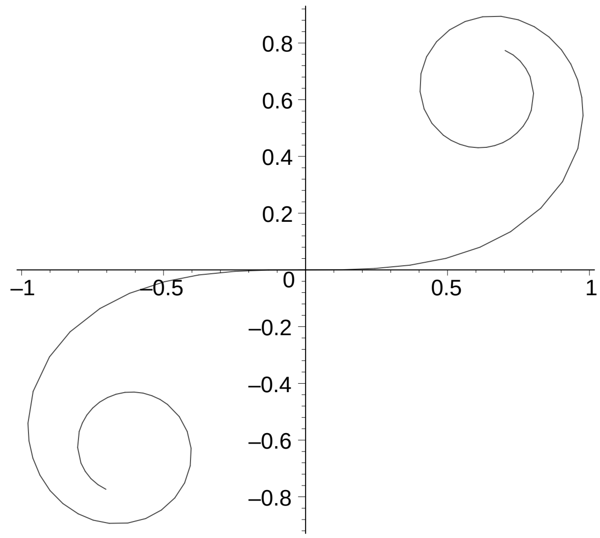


FIGURE 22. Curve with $\kappa(s) = s, \; -3 \leq s \leq 3$

4. Space Curves: Moving Frames and Torsion

4.1. Osculating Planes and Moving Frame.

4.1.1. *The osculating plane.* The curvature of a curve is defined as a measure for the rate of change of the best approximating line (the tangent line) to that curve. This is a general concept: Many geometric entities are formulated in terms of *changes* of *linear approximations*. Which linear approximations to a curve (other than tangent lines) could we investigate? For a *space* curve, it makes sense to ask for the *best approximating plane* to the curve at a given point P – the so-called *osculating plane* ω_P, cf. Fig. 23). One should think of it as *the* plane, that *almost* contains the curve segment close to P. Our aim is to make this vague idea precise – this is necessary, before we can do any calculations.

Let us look at a regular curve C and a point P on C. As usual in theoretical considerations, we assume that C given by its arc length parametrization

$$\mathbf{r}_{al} : [0, l] \to \mathbf{R}^3, \; s \mapsto \mathbf{r}_{al}(s), \text{ and } \overrightarrow{OP} = \mathbf{r}_{al}(s_0),$$

cf. Def. 2.22. Remember, that $\mathbf{t}(P) = \mathbf{r}'_{al}(s_0)$ is a (unit) parallel vector to the tangent line t_P. Furthermore, if the curvature of C at P given by $\kappa(P) = |\mathbf{r}''_{al}(s_0)|$ does *not* vanish – which we assume from now on – the principal normal vector to the curve at P is given as $\mathbf{n}(P) = \frac{\mathbf{r}''_{al}(s_0)}{|\mathbf{r}''_{al}(s_0)|}$. In particular, $\mathbf{t}(P)$ and $\mathbf{n}(P)$ are *perpendicular* and thus linearly independent. Certainly, the osculating plane ω_P should contain the tangent line t_P to the curve through P. Moreover, remember that the osculating circle of C at P was defined in Sect. 3.4 as the best approximating circle to the curve at P. The osculating circle is contained in one and only one plane in space – the plane ω_P, cf. Prop. 2.32.

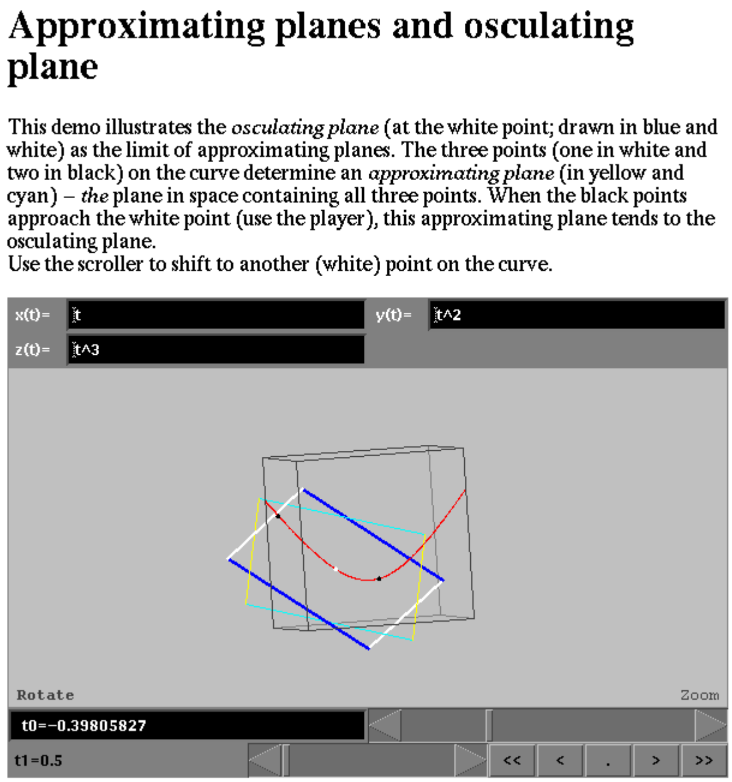


FIGURE 23. Approximating planes and osculating plane

DEFINITION 2.36. Let C be a regular space curve, and let P denote a point on C such that $\kappa(P) \neq 0$. Then the *osculating plane* ω_P of C at P is defined as the plane through P with parallel plane $sp(\mathbf{t}(P), \mathbf{n}(P))$.

REMARK 2.37.

(1) Hence, one obtains a parametrization for ω_P as
$$\omega_P = \{P_{st} | \overrightarrow{OP_{st}} = \overrightarrow{OP} + s\mathbf{t}(P) + t\mathbf{n}(P),\ s, t \in \mathbf{R}\}.$$
(2) It is obvious that ω_P contains the tangent line t_P through P – with parallel vector $\mathbf{t}(P)$.
(3) Another interpretation for ω_P is as follows: For a point Q on the curve close to P form the plane through P containing the tangent line t_P *and* the (secant) line through P and Q. When Q tends to P (along the curve), this plane tends to a *limit* plane, which is ω_P.

In general, one does not have an explicit description of the arc length parametrization for C. Nevertheless, with a curve C as in Def. 2.36 given by a *regular* parametrization $\mathbf{r} : I \to \mathbf{R}^3$ (cf. Def. 2.10) and a point P on C with $\overrightarrow{OP} = \mathbf{r}(t_0)$, it is still easy to describe a parametrization for ω_P:

LEMMA 2.38. *The osculating plane ω_P can also be described in the following ways:*

(1) $\omega_P = \{P_{st} |\ \overrightarrow{OP_{st}} = \overrightarrow{OP} + s\mathbf{r}'(t_0) + t\mathbf{r}''(t_0),\ s,t \in \mathbf{R}\}$.
(2) *ω_P is the plane through P containing the* velocity *vector* $\mathbf{v}(P) = \mathbf{r}'(t_0)$ *and the* acceleration *vector* $\mathbf{a}(P) = \mathbf{r}''(t_0)$ *for* every *parametrization of the curve C.*

PROOF:
In (2.10), we got the following decomposition of the acceleration vector: $\mathbf{a}(t_0) = \mathbf{r}''(t_0) = v'(t_0)\mathbf{t}(t_0) + v^2(t_0)\kappa(t_0)\mathbf{n}(t_0)$. Since $\mathbf{r}'(t_0)$ and the unit tangent vector $\mathbf{t}(t_0)$ are parallel, the planes $sp(\mathbf{r}'(t_0), \mathbf{r}''(t_0))$ and $sp(\mathbf{t}(P), \mathbf{n}(P))$ agree. □

EXAMPLE 2.39.
Let $\mathbf{r}$ denote the parametrization for a helix from Ex. 2.9 given as

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

Using the calculation of $\mathbf{r}$'s derivatives in Exc. 2.31 and Lemma 2.38, we obtain a parametrization for the osculating plane ω_0 at P_0 with $\overrightarrow{OP_0} = \mathbf{r}(t_0)$ by

$$\omega_0 : [a\cos t_0, a\sin t_0, bt_0] + s[-a\sin t_0, a\cos t_0, b] + t[-a\cos t_0, -a\sin t_0, 0] \quad s,t \in \mathbf{R}.$$

REMARK 2.40.
The osculating plane ω_P for C at P is the plane through P containing both the first degree and the second degree Taylor polynomium (2.3)

$$\begin{aligned} \mathbf{r}^{(1)}_{t_0}(t) &= \mathbf{r}(t_0) + (t-t_0)\mathbf{r}'(t_0); \\ \mathbf{r}^{(2)}_{t_0}(t) &= \mathbf{r}(t_0) + (t-t_0)\mathbf{r}'(t_0) + \frac{(t-t_0)^2}{2}\mathbf{r}''(t_0). \end{aligned}$$

of the vector function $\mathbf{r}$ at $t = t_0$.

4.1.2. *The moving frame.* To find a linear equation describing ω_P, one has to find a normal vector to it. Let a curve C and a point $P \in C$ be given as in Def. 2.36:

DEFINITION 2.41. The *binormal* vector $\mathbf{b}(P)$ to C at P is defined as the unit vector perpendicular to ω_P given as

$$\mathbf{b}(P) = \mathbf{t}(P) \times \mathbf{n}(P). \tag{2.24}$$

LEMMA 2.42. (1) *For a general regular parametrization* $\mathbf{r}$ *of* C *with* $\overrightarrow{OP} = \mathbf{r}(t_0)$, *one has*

$$\mathbf{b}(P) = \frac{\mathbf{r}'(t_0) \times \mathbf{r}''(t_0)}{|\mathbf{r}'(t_0) \times \mathbf{r}''(t_0)|}. \tag{2.25}$$

(2) *The vector* $\mathbf{r}'(t_0) \times \mathbf{r}''(t_0)$ *is a normal vector to* ω_P, *too.*

(3) *Each of the following three equations characterizes the points in the osculating plane:*
$Q : (x, y, z) \in \omega_P \Leftrightarrow$:
- $\mathbf{b}(P) \cdot \overrightarrow{OQ} = \mathbf{b}(P) \cdot \overrightarrow{OP}$;
- $(\mathbf{r}'(t_0) \times \mathbf{r}''(t_0)) \cdot \overrightarrow{OQ} = (\mathbf{r}'(t_0) \times \mathbf{r}''(t_0)) \cdot \overrightarrow{OP}$;
- $[\mathbf{r}'(t_0), \mathbf{r}''(t_0), \overrightarrow{OQ}] = [\mathbf{r}'(t_0), \mathbf{r}''(t_0), \overrightarrow{OP}]$.

(4) *To determine an equation for the osculating plane* ω_P *at* P *with* $\overrightarrow{OP} = \mathbf{r}(t_0)$, *proceed as follows:*
   (a) *Determine the coordinates* $[x_0, y_0, z_0] = \mathbf{r}(t_0)$ *of point* P;
   (b) *Determine the coordinates* $[a, b, c] = \mathbf{r}'(t_0) \times \mathbf{r}''(t_0)$ *of a vector perpendicular to* ω_P.
   (c) *Then* ω_P *is given by the equation* $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$; *cf. (1.22).*

PROOF:
Only 1. above has to be proved; the equations in 3. are derived as in Sect. 1.2.2. By Lemma 2.38, the planes $sp(\mathbf{r}'(t_0), \mathbf{r}''(t_0))$ and $sp(\mathbf{t}(P), \mathbf{n}(P))$ agree; hence they have the *same unit normal* vector $\mathbf{b}(P)$. □

At each point $P \in k$, the three vectors $\mathbf{t}(P), \mathbf{n}(P), \mathbf{b}(P)$ form a basis of mutually orthogonal unit vectors. Along a transit of the curve, this coordinate system $\{\mathbf{t}(P), \mathbf{n}(P), \mathbf{b}(P)\}$ will also move and twist around – it is called the *moving frame*[6] associated to the curve.

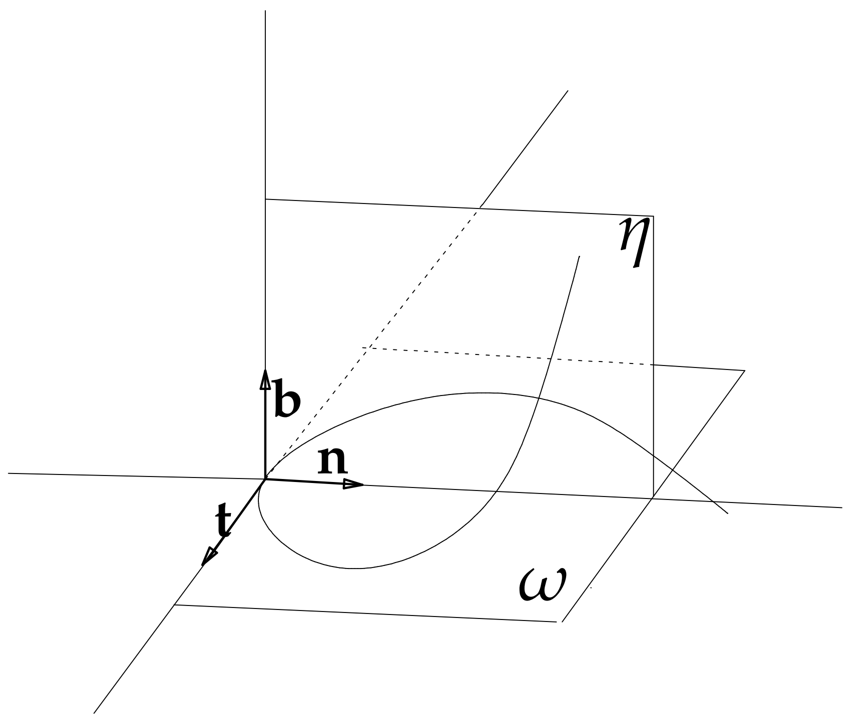


FIGURE 24. Moving frame, osculating plane and normal plane

At every point, it gives us the best perspective to view the curve, in the following sense:

- The vectors $\mathbf{t}(P)$ and $\mathbf{n}(P)$ span a parallel plane to the *osculating plane* ω_P – the best approximating plane. The vector $\mathbf{b}(P)$ is orthogonal to ω_P.
- The vectors $\mathbf{n}(P)$ and $\mathbf{b}(P)$ span a parallel plane to the *normal plane* η_P, i.e., the plane through P that is perpendicular to the tangent line t_P.

REMARK 2.43.
Given a regular parametrization $\mathbf{r}$, the determination of the moving frame at the point P with $\overrightarrow{OP} = \mathbf{r}(t_0)$ is usually performed in the following three steps:

$$1. \ \mathbf{t}(P) = \frac{\mathbf{r}'(t_0)}{|\mathbf{r}'(t_0)|};$$

[6]illustrations: Fig. 24 and Fig. 25

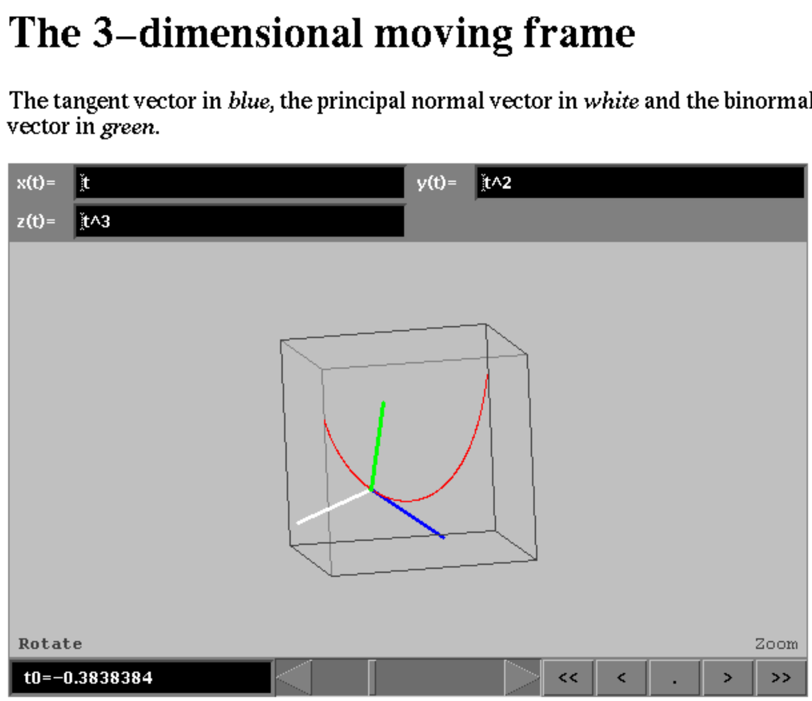


FIGURE 25. Moving frame of a space curve

$$2.\ \mathbf{b}(P) = \frac{\mathbf{r}'(t_0) \times \mathbf{r}''(t_0)}{|\mathbf{r}'(t_0) \times \mathbf{r}''(t_0)|};$$
$$3.\ \mathbf{n}(P) = \mathbf{b}(P) \times \mathbf{t}(P).$$

EXAMPLE 2.44.
Let us again look at a helix with parametrization $\mathbf{r}(t) = [a\cos t, a\sin t, bt],\ a, b > 0$. (cf. Ex. 2.9 and 2.31). Using the formulas from Rem. 2.43, we obtain at P_0 with $\overrightarrow{OP_0} = \mathbf{r}(t_0)$:

$$\begin{aligned}
\mathbf{t}(t_0) &= \frac{1}{\sqrt{a^2+b^2}}[-a\sin t_0, a\cos t_0, b]; \\
\mathbf{b}(t_0) &= \frac{1}{\sqrt{a^2+b^2}}[b\sin t_0, -b\cos t_0, a]; \qquad (2.26)\\
\mathbf{n}(t_0) &= [-\cos t_0, -\sin t_0, 0].
\end{aligned}$$

In particular, at every point P_0 along the curve, the principal normal vector $\mathbf{n}$ is parallel to the XY-plane pointing from the curve in direction of the $z-$axis. The osculating plane at P_0 is given (cf. Lemma 2.42) by the equation

$$\mathbf{b}(t_0) \cdot [x, y, z] = \mathbf{b}(t_0) \cdot \overrightarrow{OP_0}, \text{ i.e.}$$
$$b \sin t_0 \cdot x - b \cos t_0 \cdot y + a \cdot z = abt_0.$$

4.2. Curvature, torsion, and Frenet's equations. In this section, we shall give a precise description of how the moving frame moves and twists when we move along a given curve. Again, let us suppose, that a space curve C is given by its arc length parametrization $\mathbf{r}_{al} : I \to \mathbf{R}^3,\ s \mapsto \mathbf{r}_{al}(s)$, cf. Def. 2.22, and $\overrightarrow{OP} = \mathbf{r}_{al}(s_0)$. Moreover, we suppose that the curvature does *not* vanish at P, i.e., $\kappa(P) \neq 0$. Remember, that the curvature $\kappa(P)$ was defined as a measure for the rate of change of the unit vector field $\mathbf{t}$ in a neighbourhood of P. This is expressed in formula (2.9):

$$\mathbf{t}'(P) = \kappa(P)\mathbf{n}(P).$$

To express the change of the whole moving frame $\{\mathbf{t}(P), \mathbf{n}(P), \mathbf{b}(P)\}$, one would need similar formulas for the rate of change of the t*wo other* moving base vectors. It will turn out, that it is enough to have an expression for $\mathbf{b}'(P)$ – which we are going to develop now:

Let us first give a *geometric interpretation* of the entity *torsion*[7] to be defined: The binormal vector $\mathbf{b}(P)$ is by definition a unit vector orthogonal to the osculating plane ω_P. Hence, the rate of change of $\mathbf{b}(P)$ is equivalent to the rate of change of ω_P; hence it measures, how quickly the osculating plane changes when moving away from P along the curve, or equivalently, how quickly the curve disappears or twists away from ω_P.

The following lemma is needed in the definition of torsion:

LEMMA 2.45. *At all points of the curve C, the vector $\mathbf{b}'(P)$ is a multiple of the principal normal vector $\mathbf{n}(P)$.*

PROOF:
It is enough to prove, that $\mathbf{b}'(P)$ is orthogonal to both $\mathbf{t}(P)$ and $\mathbf{b}(P)$. This is a consequence of the fundamental trick (Prop. 2.5) and of the definition of curvature (Def. 2.28):

- $\mathbf{t}(s) \cdot \mathbf{b}(s) = 0 \Rightarrow \mathbf{t}'(s) \cdot \mathbf{b}(s) + \mathbf{t}(s) \cdot \mathbf{b}'(s) = 0 \Rightarrow$
  $\mathbf{t}(s) \cdot \mathbf{b}'(s) = -(\mathbf{t}'(s) \cdot \mathbf{b}(s)) = -\kappa(s)(\mathbf{n}(s) \cdot \mathbf{b}(s)) = 0.$
- $\mathbf{b}(s) \cdot \mathbf{b}(s) = 1 \Rightarrow \mathbf{b}'(s) \cdot \mathbf{b}(s) = 0.$

□

[7] illustration: opposite page

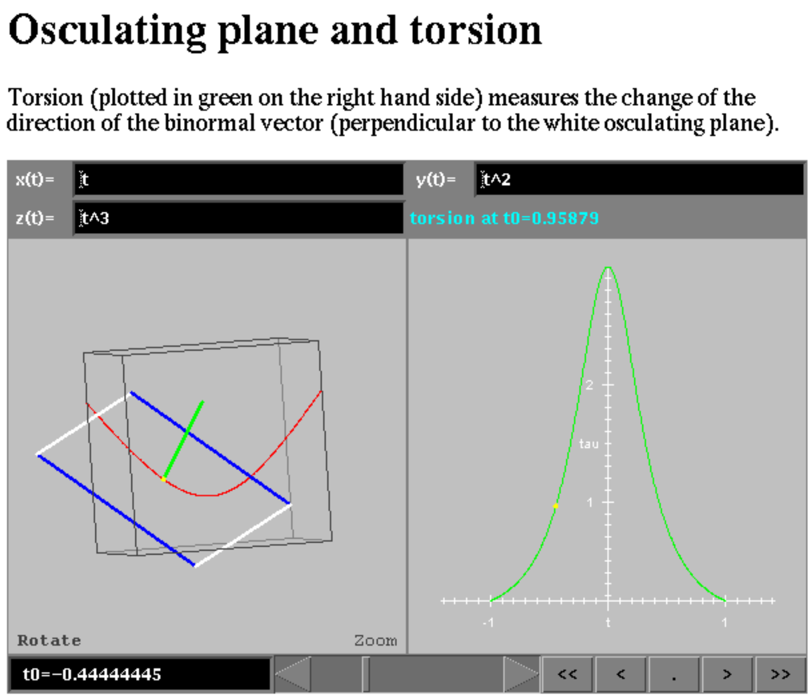


FIGURE 26. Osculating plane and torsion

4.2.1. *Definition of the torsion of a space curve.*

DEFINITION 2.46. Let C be a regular curve, and $P \in C$ a point on C with $\overrightarrow{OP} = \mathbf{r}_{al}(s)$ such that $\kappa(P) \neq 0$. Then the *torsion* $\tau(P) = \tau(s)$ of the curve C at P is defined to be the real number satisfying

$$\mathbf{b}'(s) = -\tau(s)\mathbf{n}(s). \tag{2.27}$$

REMARK 2.47.
The curvature of a space curve at a point was defined to be non-negative; its torsion can be both *positive, negative or zero*. The sign in Def. 2.46 is somehow arbitrary, and some authors write $\tau(s)$ instead of our $-\tau(s)$.

Curvature and torsion together determine how the whole moving frame develops; this is contained in the following result known as *Frenet's equations:*

THEOREM 2.48. [a] *(Frenet) Let C be a regular curve with moving frame* $\{\mathbf{t}, \mathbf{n}, \mathbf{b}\}$ *as functions of the arc length parameter s. Then,*

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

[a]Here is a biography of the French mathematician astronomer and meteorologist Jean Frenet. The equations should rightly be called Frenet-Serret formulae, to honour moreover the French mathematician and physicist Joseph Serret.

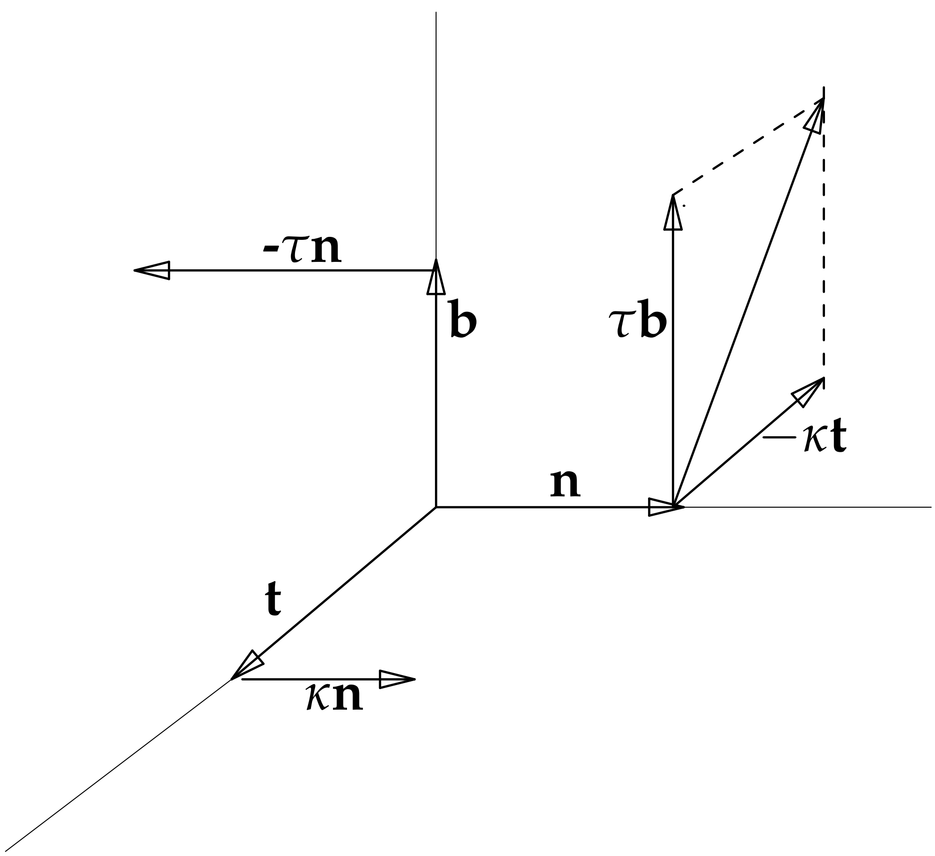


FIGURE 27. Frenet's equations

PROOF:
The first and third equation are just the defining equations for curvature and torsion

of a space curve. The second equation is a consequence of the two others and of the fundamental trick (Prop. 2.5). We calculate the projections cf. (3.1.1) of the vector $\mathbf{n}'(s)$ on the 3 coordinate vectors $\mathbf{t}(s)$, $\mathbf{n}(s)$, and $\mathbf{b}(s)$:

- $\mathbf{t}(s) \cdot \mathbf{n}(s) = 0 \Rightarrow \mathbf{n}'(s) \cdot \mathbf{t}(s) = -(\mathbf{t}'(s) \cdot \mathbf{n}(s)) = -\kappa(s)$;
- $\mathbf{n}(s) \cdot \mathbf{n}(s) = 1 \Rightarrow \mathbf{n}'(s) \cdot \mathbf{n}(s) = 0$;
- $\mathbf{n}(s) \cdot \mathbf{b}(s) = 0 \Rightarrow \mathbf{n}'(s) \cdot \mathbf{b}(s) = -(\mathbf{n}(s) \cdot \mathbf{b}'(s)) = \tau(s)$.

□

4.2.2. *Curvature and torsion determine space curves.* For space curves, there is a result analogous to the one described in Sect. 3.5:

PROPOSITION 2.49. *Given two differentiable functions $\kappa, \tau : [0, l] \to \mathbf{R}$ with $\kappa(s) > 0$ for all $0 < s < l$. There is a curve C parameterised by arc length $\mathbf{r}_{al}$ with curvature function κ and torsion function τ. This curve is* uniquely determined *up to a* rigid motion *in space, i.e., a combination of a* translation *and a* rotation.

We will not give a proof of Prop. 2.49: It relies on the fact, that the system of differential equations given by Frenet's equations has a unique solution for given initial conditions.

4.2.3. *Calculation of the torsion.* Frenet's equations (Thm. 2.48) 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 2.50. *Let C be a curve with regular parametrization, and let P be a point on C with $\overrightarrow{OP} = \mathbf{r}(t)$ and $\kappa(P) \neq 0$. Then,*

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

PROOF:
We reuse the formulas (2.10) for the velocity and the acceleration vector in the following form:

$$\begin{aligned} \mathbf{r}'(t) &= v(t)\mathbf{t}(s(t)); \\ \mathbf{r}''(t) &= v'(t)\mathbf{t}(s(t)) + v^2(t)\kappa(s(t))\mathbf{n}(s(t)). \end{aligned} \tag{2.28}$$

A calculation of their cross product yields:

$$\mathbf{r}'(t) \times \mathbf{r}''(t) = v^3(t)\kappa(s(t))\mathbf{b}(s(t)). \tag{2.29}$$

(Why?). Thus,

$$[\mathbf{r}'(t), \mathbf{r}''(t), \mathbf{r}'''(t)] = (\mathbf{r}'(t) \times \mathbf{r}''(t)) \cdot \mathbf{r}'''(t) = v^3(t)\kappa(s(t))(\mathbf{b}(s(t)) \cdot \mathbf{r}'''(t)). \tag{2.30}$$

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. Lemma 2.38. Hence, we may use first the fundamental trick (Lemma 2.5) and then the last of Frenet's equations, cf. Thm. 2.48, to obtain:

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

Applying the second equation in (2.28) once again, we derive:

$$\mathbf{b}(s(t)) \cdot \mathbf{r}'''(t) = v^3(t)\kappa(s(t))\tau(s(t)). \tag{2.31}$$

Substituting (2.31) into (2.30), we obtain:

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

On the other hand, (2.29) tells us, that $|\mathbf{r}'(t) \times \mathbf{r}''(t)|^2 = v^6(t)\kappa^2(s(t))$, which implies the formula in Prop. 2.50 for the torsion $\tau(P) = \tau(s(t))$. □

REMARK 2.51.
From (2.31), we deduct, that the torsion $\tau(P)$ has the same sign as the entity $\mathbf{b}(P) \cdot \mathbf{r}_{al}'''(s)$ – 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 opposite 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. 4.4.

EXAMPLE 2.52.
For the helix from Ex. 2.9 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. 2.31.3:

$$\begin{aligned} \mathbf{r}'''(t) &= [a\sin t, -a\cos t, 0]; \\ \mathbf{r}'(t) \times \mathbf{r}''(t) &= [ab\sin t, -ab\cos t, a^2]; \\ (\mathbf{r}'(t) \times \mathbf{r}''(t)) \cdot \mathbf{r}'''(t) &= a^2 b; \\ \tau(P) &= \frac{b}{a^2+b^2}. \qquad (2.32) \end{aligned}$$

Note, that both curvature (cf. (2.20)) and torsion are *constant* along the helix.

4.3. What happens to curvature under projection of a space curve to a plane? Since we usually only can draw plane curves (on a sheet of paper or a computer screen), the following question arises naturally:

Let C denote a *space* curve and α a plane in space. Projecting each point of C orthogonally onto α produces a *plane* curve C_α, cf. Fig. 28. What relation is there between the curvature functions associated to C and to C_α? Which plane $\alpha \subset \mathbf{R}^3$ should one choose in order to preserve curvature features in the best possible way?

To formulate the answer, we fix a point $P \in C$ and its projection P_α on the projected curve C_α. At P, the curve C has a *unit tangent vector* $\mathbf{t}_P$ and an *osculating plane* ω_P. Then the answer to the question above can be given in terms of two *angles* (cf.Sect. 1.3.2.2 and Fig. 28): Let

- ϕ denote the angle between the osculating plane ω_P and the projection plane α, and
- θ denote the angle between the tangent vector $\mathbf{t}_P$ and the projection plane α.

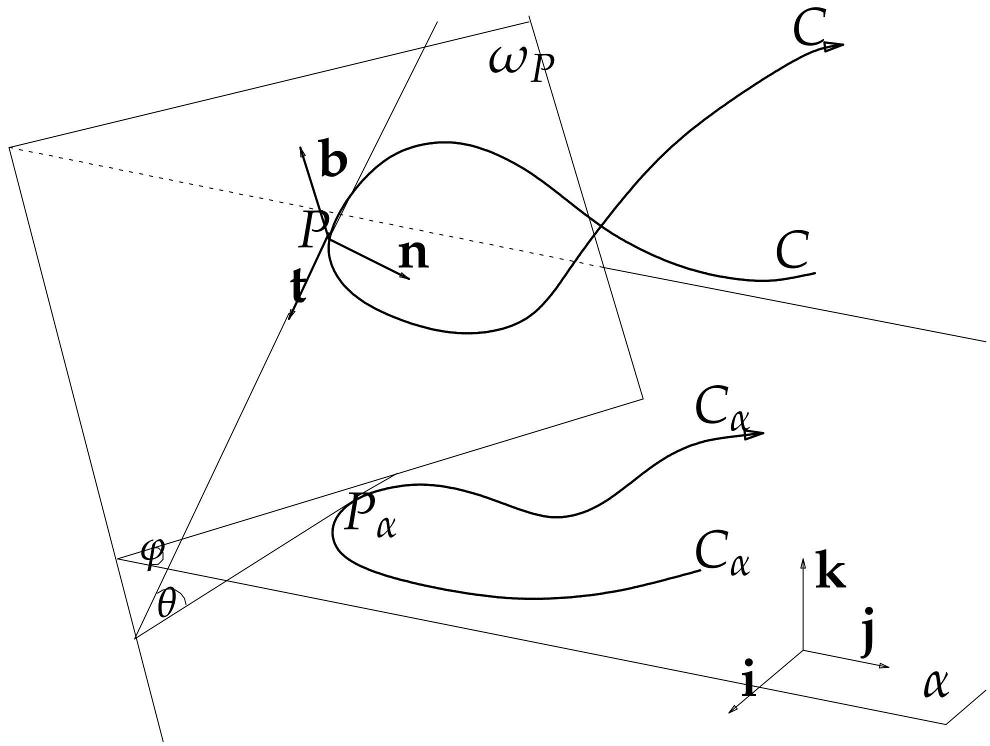


FIGURE 28. A space curve with a plane projection

PROPOSITION 2.53. *The curvature of the projected curve C_α at P_α is related to the curvature of C at P in the following way:*

$$|\kappa(P_\alpha)| = \kappa(P) \cdot \frac{\cos\phi}{(\cos\theta)^3}.$$

COROLLARY 2.54. *If C is projected onto the osculating plane $\alpha = \omega_P$, then $|\kappa(P_\alpha)| = \kappa(P)$, i.e., the numerical values of the curvatures agree.*

PROOF:
For $\alpha = \omega_P$, both angles $\phi = \theta = 0$. □

This means, that the osculating plane at a point P is a good choice for viewing the curve in the neighboorhood of P. The curvatures of the original and the projected curve agree at P, and they will not differ very much in the vicinity of P.

PROOF OF PROP. 2.53:
We start with the following trick that makes notation and calculation much easier: We perform a rigid motion (a combination of a rotation and a translation) that sends α into the XY-plane. Remark that a rigid motion does not affect curvature! From now on, we assume that α is in fact the XY-plane with normal vector $\mathbf{n} = \mathbf{k}$.

Let the curve C be given by a parametrization $\mathbf{r} : I \to \mathbf{R}^3,\ \mathbf{r}(t) = [x(t), y(t), z(t)]$, such that $\overrightarrow{OP} = \mathbf{r}(t_0)$. Its projection onto the XY-plane can thus be given the parametrization $\mathbf{r}_0 : I \to \mathbf{R}^2,\ \mathbf{r}_0(t) = [x(t), y(t)]$ with $\overrightarrow{OP_\alpha} = \mathbf{r}_0(t_0)$. Let us compare first the denominators and then the numerators for the curvature formulas (2.13) and (2.15), all to be taken at t_0):

$$\mathbf{r}' \times \mathbf{r}'' = [y'z'' - y''z', z'x'' - z''x', x'y'' - x''y'];$$
$$[\mathbf{r}_0', \mathbf{r}_0''] = x'y'' - x''y' = \mathbf{k} \cdot (\mathbf{r}' \times \mathbf{r}'') = |(\mathbf{r}' \times \mathbf{r}'')| \cos\phi.$$

The last equation uses Prop. 1.5 and the fact that the angle between $\mathbf{k}$ and $\mathbf{r}' \times \mathbf{r}''$, is numerically equal to the angle ϕ between their normal planes (cf. Sect. 1.3.2.2), i.e, the angle between α (the XY-plane) and the osculating plane ω_P.

Looking at denominators, $\mathbf{r}_0' = [x', y']$ is the projection of $\mathbf{r}' = [x', y', z']$ onto α (the XY-plane), and hence (cf. Sect. 1.3.2.2), $|\mathbf{r}_0'| = |\mathbf{r}'| \cos\theta$. Combining both results, we obtain:

$$|\kappa(P_\alpha)| = \frac{[\mathbf{r}_0', \mathbf{r}_0'']}{|\mathbf{r}_0'|^3} = \frac{|(\mathbf{r}' \times \mathbf{r}'')| \cos\phi}{|\mathbf{r}'|^3 (\cos\theta)^3} = \kappa(P) \frac{\cos\phi}{(\cos\theta)^3}.$$

□

4.4. The local canonical form of a 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 parameterized by arc length $\mathbf{r}_{al} : I \to \mathbf{R}^3$ 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 (cf. (2.2))[8] at O:

$$\mathbf{r}_0^{(3)}(s) = \mathbf{r}(0) + s\mathbf{r}'(0) + \frac{s^2}{2}\mathbf{r}''(0) + \frac{s^3}{6}\mathbf{r}'''(0). \tag{2.33}$$

By our assumptions and using Frenet's 2nd equation from Thm. 2.48 we obtain:

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

Substituting this into (2.33), we get for the 3rd degree Taylor polynomial (2.2):

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

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 (being always non-negative anyhow), we obtain as a consequence (compare Rem. 2.51):

COROLLARY 2.55. *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 figures on next page show 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:

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

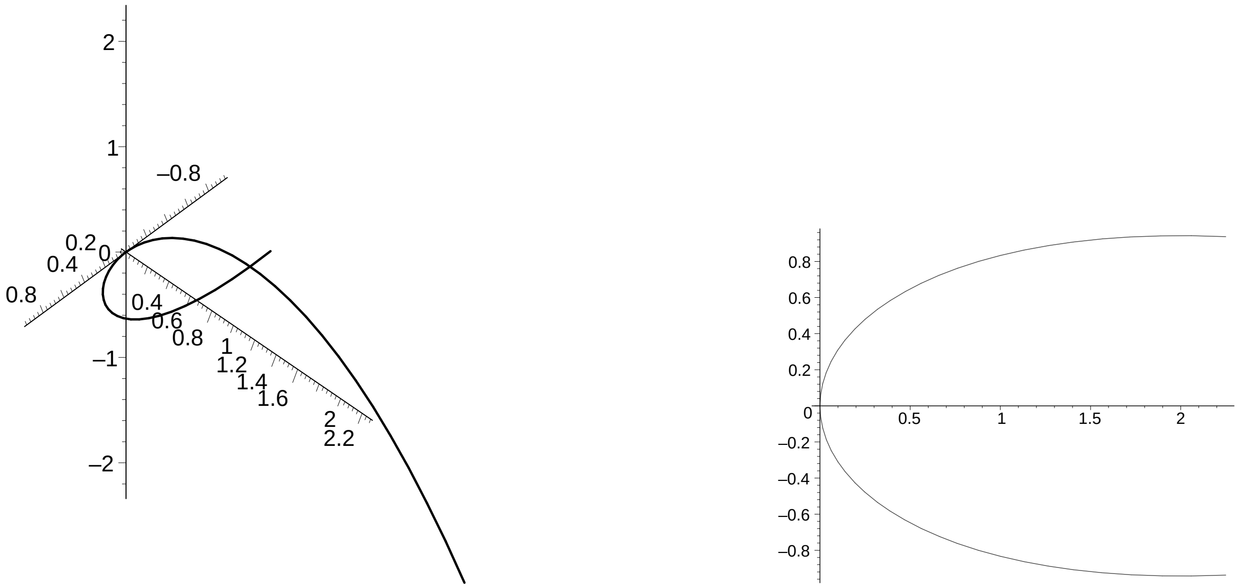


FIGURE 29. A space curve and its projection to the **nt**-plane

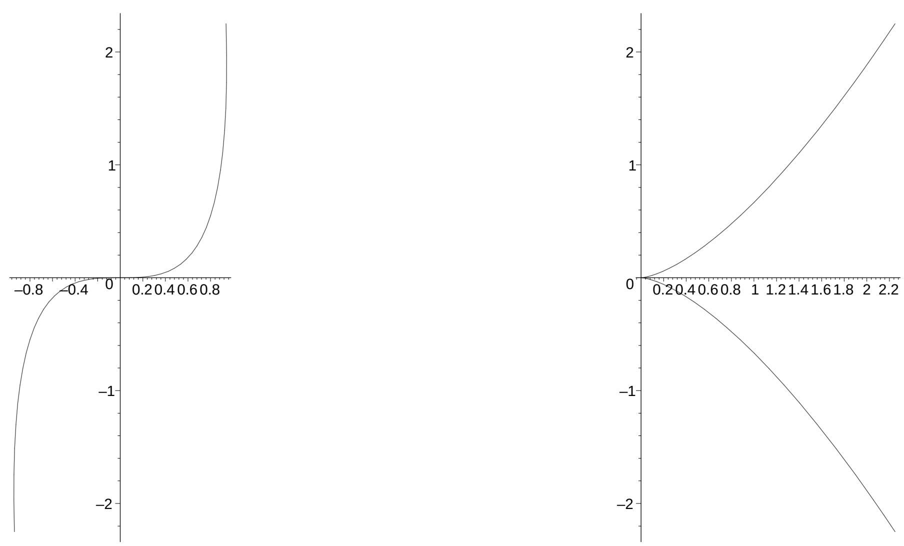


FIGURE 30. Projections to the **tb**-plane and to the **nb**-plane

5. How to use the geometric laboratory - an example

Suppose you want to know more about the osculating circle to a curve at various points and how it arises as a limit circle of approximating circles to that curve. This is explained in the applet Approximating circles and osculating circles, which you can access by clicking on the framed area – provided your browser supports Java2.

You will find a standard initial curve with parametrization $[x(t) = t, y(t) = t^2]$, $-1 \leq t \leq 1$.

Suppose you want to investigate a different cuve, say the curve with parametrization $[x(t), y(t)] = [\cos(2t), \sin(3t)]$, $\frac{\pi}{2} \leq t \leq \frac{3}{2}\pi$. You ”overwrite” t in the black field to the right of $x(t) =$ by $\cos(2t)$ – *RETURN* – and also t^2 in the field to the right of $y(t) =$ by $\sin(3t)$ – *RETURN*. Now click on the field marked $t0 = -1.0$ below the graphics. Replace -1.0 in the field to the right of *From:* by 1.57, click on *From:* and replace 1.0 in the field right of *To:* by 4.72. Use the scroller (to the right of the field marked $t0 =$) to move the points to a more interesting position (e.g., $t0 = 2.64$) You will obtain what you see in Fig. 31, with the red curve, a blue approximating circle which is the circle through the white and the two black points and the approximating circle through the white point.

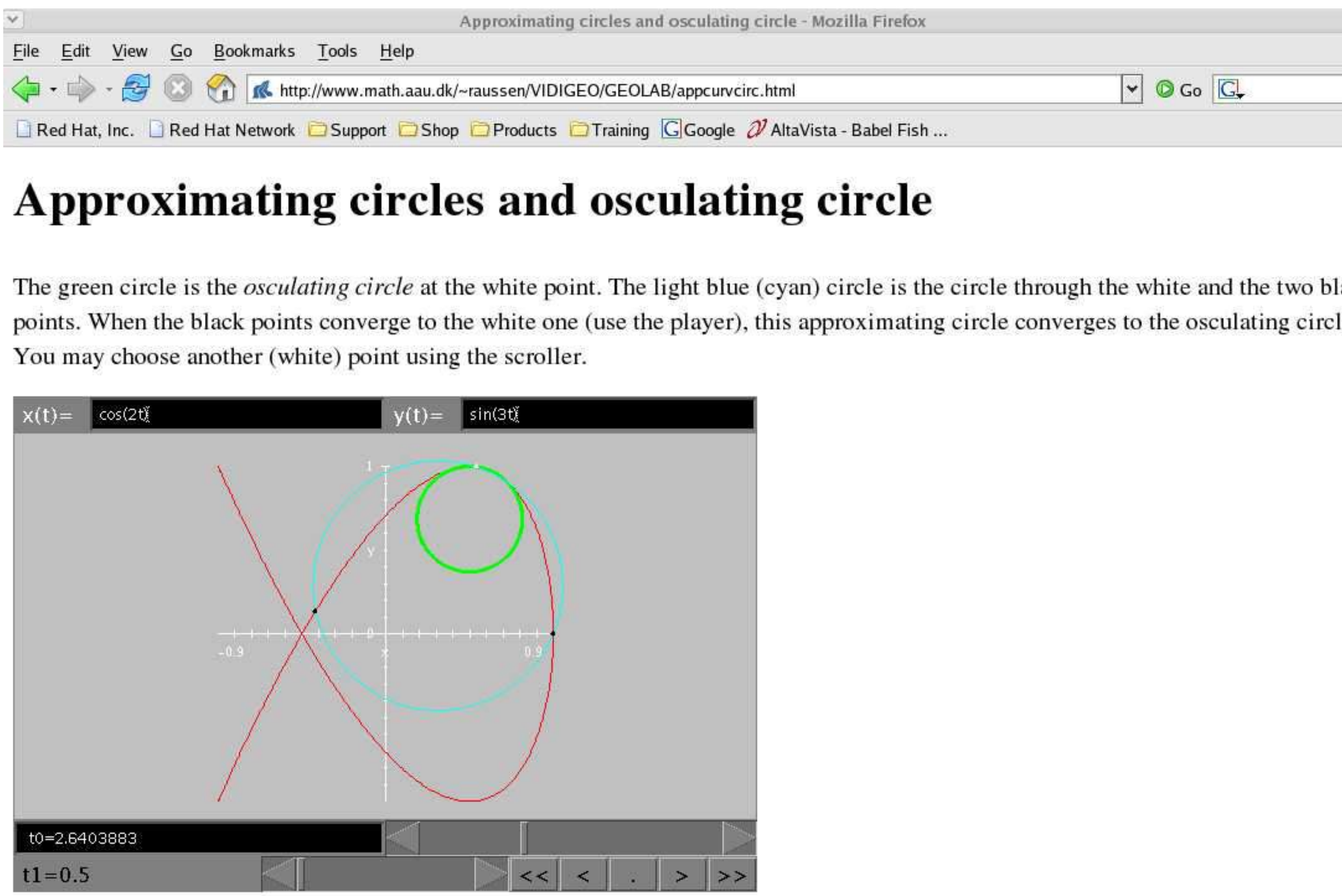


FIGURE 31. Applet screen

To observe how the approximating circle approaches the osculating circle as the black points converge to the white point, you should use the player (to the left of the region marked $t1 = 0.5$). Click on the right arrow – or double arrow, if you are less patient and watch what happens. If you want to see approximating and osculating circles at others than the given initial white point, you can scroll to another position $t0 =$.

CHAPTER 3

Regular Surfaces

1. Parametrizations of surfaces

Examples of *surfaces* abound in everyday life: Balloons, tubes, cans, soap films, the surface of our planet earth are all physical models of surfaces. In order to study the geometry of these objects, one needs *coordinates* to make calculations. Of course, all these surfaces can be thought of as embedded in Euclidean space $\mathbf{E}^3$ – which means 3 coordinates. But just as a (space) curve needs only 1 coordinate, the very definition of a *surface* is that it can be described (in a nice way) using just *two* coordinates:

- We are used to describe points on the surface of the earth by the *two* geographical coordinates: *longitude* and *latitude*.
- One can (and does!) produce cylindrical cans by rolling a plane (2-dimensional) piece of metal.

1.1. Vector functions of two variables and partial differentiation.

1.1.1. *Smooth vector functions and partial derivatives.*

DEFINITION 3.1. Let $D \subset \mathbf{R}^2$ denote an open subset.

- A function $f : D \to \mathbf{R}$ of two variables is called *smooth* (C^∞), if *all* (higher) partial derivatives exist, and, moreover, are continuous functions.
- A vector function $\mathbf{r} : D \to \mathbf{R}^3$, $\mathbf{r}(u,v) = [x(u,v), y(u,v), z(u,v)]$ of two variables is called *smooth* (C^∞), if its coordinate functions $x, y, z : D \to \mathbf{R}$ are all smooth.
- The partial derivatives of a smooth function $f : D \to \mathbf{R}$ with respect to the variables u and v at a point $(u_0, v_0) \in D$ are denoted $f_u(u_0, v_0),\ f_v(u_0, v_0) \in \mathbf{R}$. Likewise, the partial derivatives of a vector function $\mathbf{r} : D \to \mathbf{R}^3$ at $(u_0, v_0) \in D$ are the vectors

$$\mathbf{r}_u(u_0,v_0) = [x_u(u_0,v_0), y_u(u_0,v_0), z_u(u_0,v_0)] \in \mathbf{R}^3, \text{ and}$$
$$\mathbf{r}_v(u_0,v_0) = [x_v(u_0,v_0), y_v(u_0,v_0), z_v(u_0,v_0)] \in \mathbf{R}^3.$$

EXAMPLE 3.2.

(1) The function $f : \mathbf{R}^2 \to \mathbf{R}$, $f(u,v) = u^2 + u\sin v - e^v$ is smooth. Its partial derivatives are given as
$$f_u(u,v) = 2u + \sin v, \; f_v(u,v) = u\cos v - e^v.$$

(2) The vector function $\mathbf{r} : \mathbf{R}^2 \to \mathbf{R}^3$, $\mathbf{r}(u,v) = [u^2 - v^2, uv, \sin u \cos v]$ is smooth with partial derivatives
$$\mathbf{r}_u(u,v) = [2u, v, \cos u \cos v], \; \mathbf{r}_v(u,v) = [-2v, u, -\sin u \sin v].$$

1.1.2. *The chain rule.* In the sequel, it will be important to be able to differentiate composite vector functions. In particular, we will look at composite functions $\mathbf{x}(t) = \mathbf{r}(u(t), v(t))$ representing particular space curves. The following result should be known from basic courses as a particular case of the *chain rule*:

PROPOSITION 3.3. *Let $D \subset \mathbf{R}^2$ denote an open subset, let $\mathbf{r} : D \to \mathbf{R}^3$ denote a smooth vector function, and let $u, v : I \to \mathbf{R}$ denote smooth functions such that $(u(t), v(t)) \in D$ for all t in the interval I. Then, the composite function $\mathbf{x} : I \to \mathbf{R}^3$ defined as $\mathbf{x}(t) = \mathbf{r}(u(t), v(t))$ is smooth and has the derivative*

$$\mathbf{x}'(t) = u'(t)\mathbf{r}_u(u(t), v(t)) + v'(t)\mathbf{r}_v(u(t), v(t)). \tag{3.1}$$

1.2. Coordinate patches. Intuitively, we wish to provide a surface S with a (curved) coordinate system with *two* coordinates - just as the coordinate grid consisting of meridians and parallel circles on a spherical surface. A technical formulation is given in terms of a (vector) map "projecting" the rectangular coordinate grid onto (a subset of) the plane to a curved coordinate system on S. More precisely, we require:

DEFINITION 3.4. Let $D \subset \mathbf{R}^2$ denote an open subset. A *smooth* vector function $\mathbf{r} : D \to \mathbf{R}^3$ of two variables is called a *parametrization* (or *coordinate patch*) for the surface $S \subset \mathbf{E}^3$ consisting of all points P with $\overrightarrow{OP} = \mathbf{r}(u, v)$ with $(u, v) \in D$ if

(1) $\mathbf{r}$ is a one-to-one (injective) map (i.e., every point in S corresponds to a *unique* point in D);

(2) The partial derivatives

$$\mathbf{r}_u(u_0, v_0) = [x_u(u_0, v_0), y_u(u_0, v_0), z_u(u_0, v_0)] \text{ and}$$

$$\mathbf{r}_v(u_0, v_0) = [x_v(u_0, v_0), y_v(u_0, v_0), z_v(u_0, v_0)]$$

are *linearly independent* at *every* point $(u_0, v_0) \in D$.

A subset $S \subset \mathbf{R}^3$ that has a coordinate patch $\mathbf{r}$ as above, is called a *regular surface*.

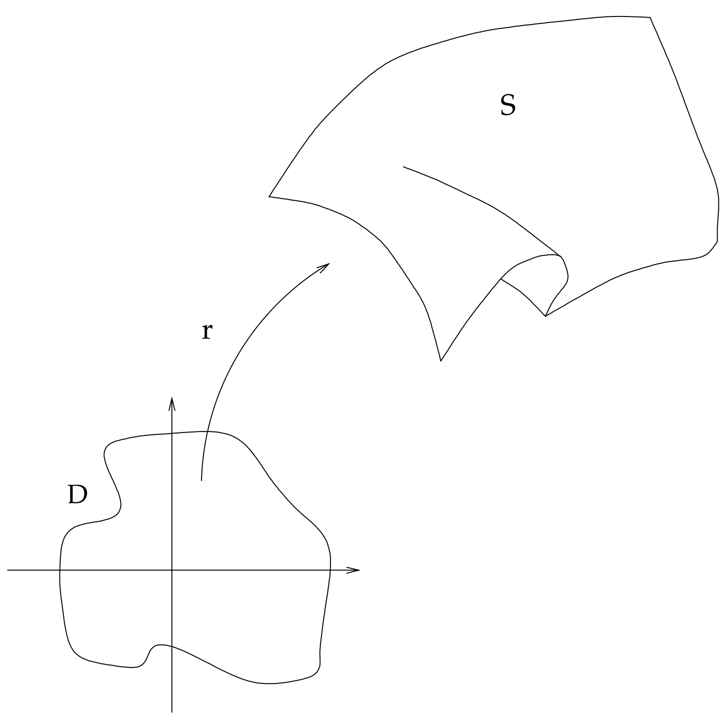


FIGURE 1. Parametrization of a surface

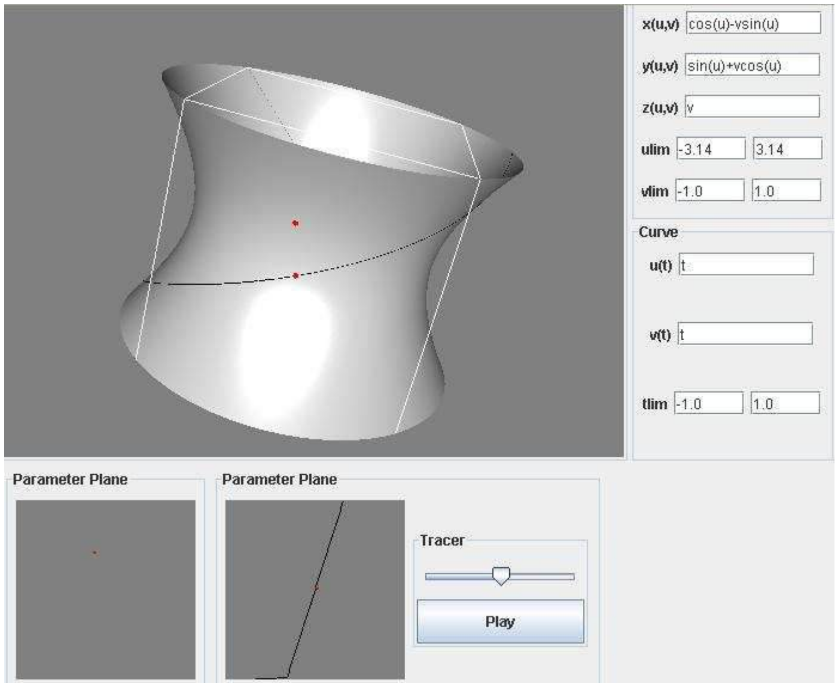


FIGURE 2. Illustration of a parametrization from the geometric laboratory

REMARK 3.5.

(1) Obviously, the first condition is needed in order to give points on the surface a *unique* set of coordinates – in D. The second condition is more intricate: it is necessary to ensure the existence of tangent planes (cf. Def. 3.10) at every point of the surface.
(2) In general, it is not possible to cover a surface with just *one* coordinate patch – this happens already for the sphere. In that case, one covers the whole surface by two or more coordinate patches.

EXAMPLE 3.6.

(1) For fixed values $\rho, a > 0$, the vector function
$$\mathbf{r} :]-a, a[\times]0, 2\pi[\to \mathbf{R}^3, \ \mathbf{r}(u, v) = [\rho \cos v, \rho \sin v, u]$$
is a parametrization for a *cylinder* with radius ρ and the Z-axis as axis. Here, we use *cylindrical coordinates*.
(2) For a fixed value $\rho > 0$, the vector function
$$\mathbf{r} :]0, \pi[\times]0, 2\pi[\to \mathbf{R}^3, \ \mathbf{r}(u, v) = [\rho \sin u \cos v, \rho \sin u \sin v, \rho \cos u]$$

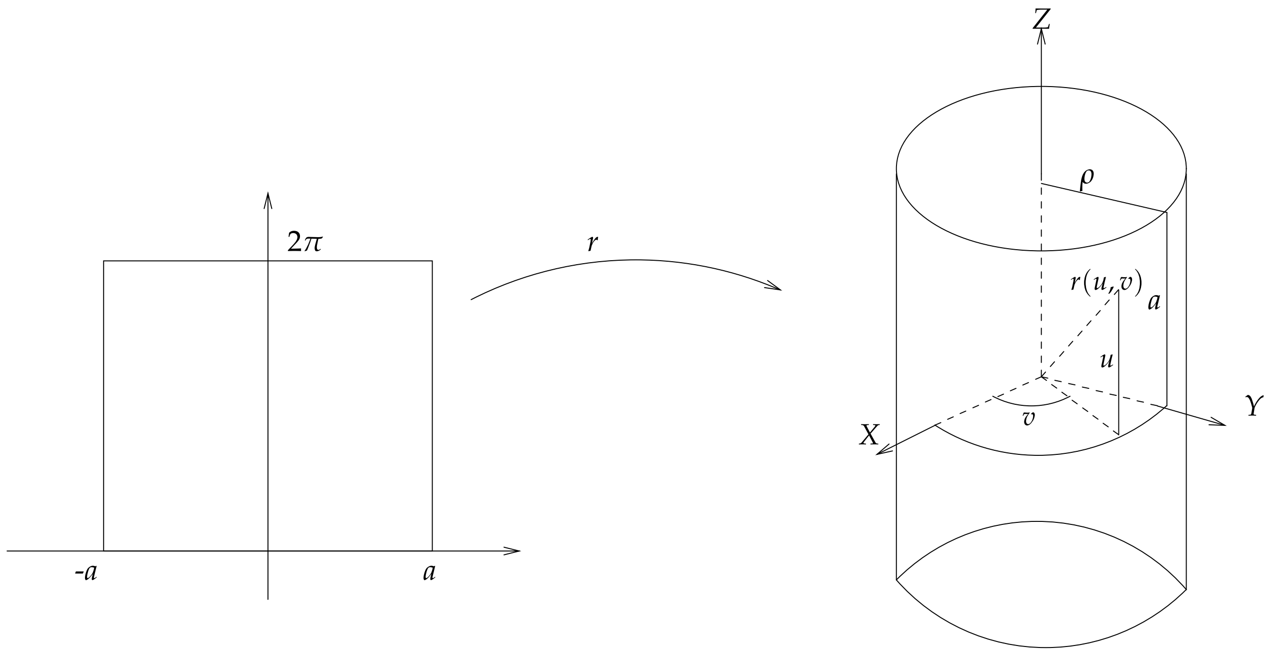


FIGURE 3. Parametrization of a cylinder

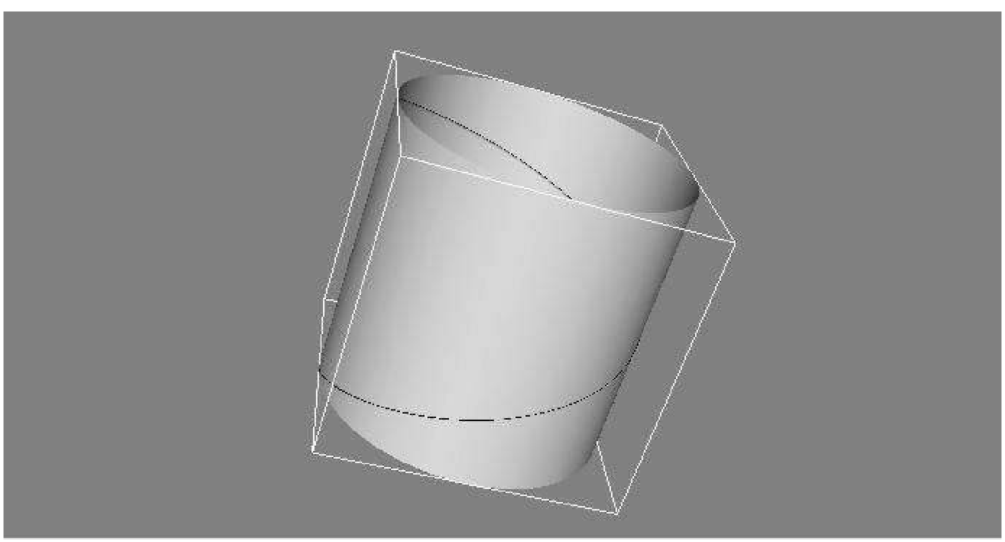


FIGURE 4. Cylinder surface (GEOLAB)

is a parametrization for the *sphere* with radius ρ with center at the origin. It represents points on the sphere in *spherical coordinates* corresponding to latitude and longitude.

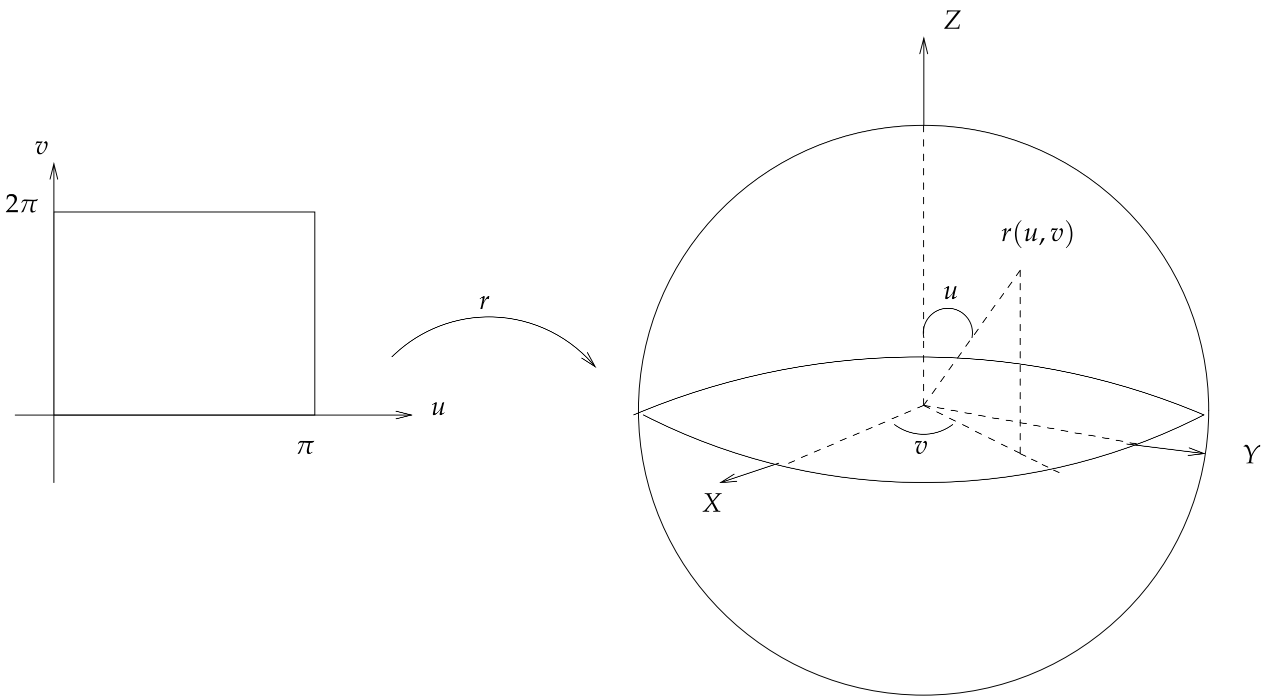


FIGURE 5. Parametrization of a sphere

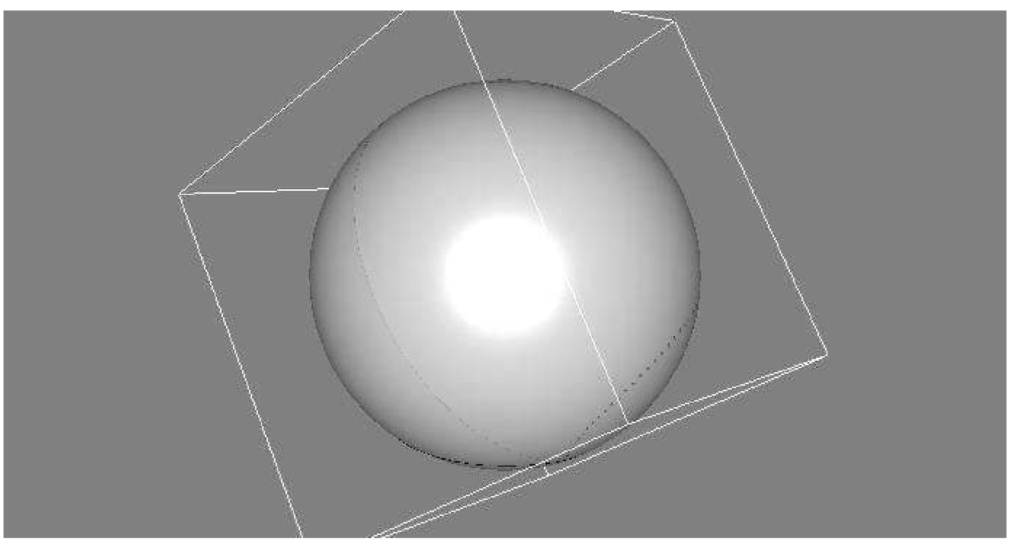


FIGURE 6. Sphere surface (GEOLAB)

Remark that this parametrization does *not* cover the *poles*, which correspond to $u = 0$, resp. $u = \pi$. Both conditions for a parametrization fail for these values of u. But the sphere is a surface at these points, too; you may obtain a parametrization for the polar regions by combining the parametrization above with a rotation of the sphere around an axis *not* containing the poles. A similar remark applies to the 0-meridian (the "Greenwich meridian").

(3) Let $f : D \to \mathbf{R}$ denote a smooth function. The *graph surface* of f consists of all points with coordinates $[u, v, f(u, v)] \in \mathbf{R}^3$. It is a surface with parametrization $\mathbf{r} : D \to \mathbf{R}^3,\ \mathbf{r}(u, v) = [u, v, f(u, v)]$.

In particular, a plane given by the equation $z = ax + by$ is the graph of $f_1 : \mathbf{R}^2 \to \mathbf{R},\ f_1(x, y) = ax + by$; the graph of $f_2 : \mathbf{R}^2 \to \mathbf{R},\ f_2(x, y) = x^2 + y^2$ is a parabolic surface of revolution obtained by rotating a parabola around the Z-axis.

(4) The vector function

$$\mathbf{r} :] - a, a[\times]0, 2\pi[\to \mathbf{R}^3,\ \mathbf{r}(u, v) = [u \cos v, u \sin v, u]$$

is a parametrization of a double *cone*. Its intersection with the horizontal plane $z = u$ is a circle with radius $|u|$ – but for $u = 0$, this circle degenerates to a single point! Conclusion: This cone does *not* satisfy the requirements of Def. 3.4, unless the line $u = 0$ in the domain of $\mathbf{r}$ corresponding to the "top point" $P : [0, 0, 0]$ in the cone is removed.

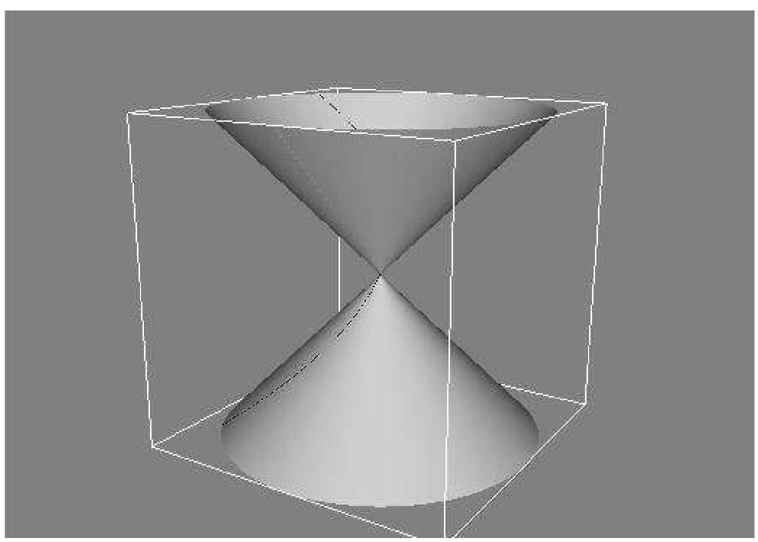


FIGURE 7. Double cone (GEOLAB)

(5) The vector function

$$\mathbf{r} :]0, 2\pi[\times]0, a[,\ \mathbf{r}(u, v) = [v \cos u, v \sin u, u]$$

is a parametrization of a *helicoid*. You may verify that this parametrization satisfies the requirements of Def. 3.4. Look at Fig. 9.

How does the definition of a parametrization correspond to the curved coordinate system? Let $\mathbf{r} : D \to \mathbf{R}^3$ denote a coordinate patch for S.

DEFINITION 3.7. The *first parameter curve* through $P_0 \in S$ with $\overrightarrow{OP_0} = \mathbf{r}(u_0, v_0)$ arises from parametrization $\mathbf{r}(u, v_0)$ looked upon as a vector function of the *single* variable u (with u in an interval containing u_0). It consists of the points $P_{(u,v_0)}$ with $\overrightarrow{OP_{(u,v_0)}} = \mathbf{r}(u, v_0)$ in the image of the *parallel* to the u-axis through (u_0, v_0).
Similarly, the *second parameter curve* through P_0 arises from the parametrization $\mathbf{r}(u_0, v)$ as a vector function of the single variable v.

By definition, the partial derivative $\mathbf{r}_u(u_0, v_0)$ is the *velocity* vector of the first parameter curve at P_0 – and thus tangent to that curve. Likewise, $\mathbf{r}_v(u_0, v_0)$ is tangent to the second parameter curve at P_0.

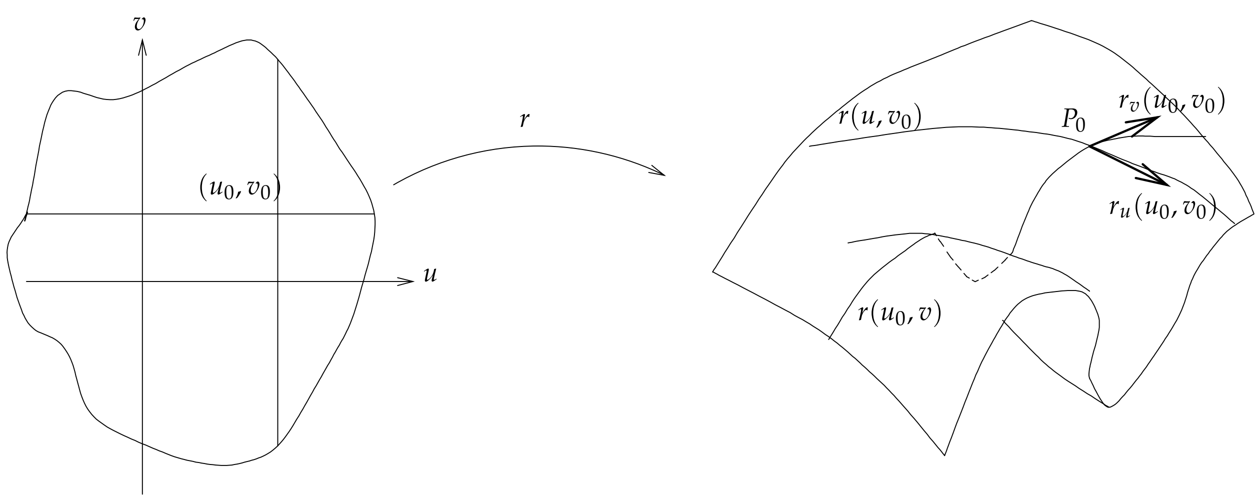


FIGURE 8. Parameter curves with tangent vectors

EXAMPLE 3.8.

(1) The first parameter curves for the cylinder with parametrization $\mathbf{r} :]-a, a[\times]0, 2\pi[\to \mathbf{R}^3$, $\mathbf{r}(u,v) = [\rho\cos v, \rho\sin v, u]$ from Ex. 3.6.1 are *vertical line segments* on the cylinder; the second parameter curves are *horizontal circles*. We calculate tangent vectors to the parameter curves as
$$\mathbf{r}_u(u_0,v_0) = [0,0,1];\ \mathbf{r}_v(u_0,v_0) = [-\rho\sin v_0, \rho\cos v_0, 0].$$
(2) The helicoid with parametrization
$$\mathbf{r} :]0,2\pi[\times]0,a[,\ \mathbf{r}(u,v) = [v\cos u, v\sin u, u]$$
has a system of *helices* as 1. parameter curves, whereas the 2. parameter curves are *straight line segments*!

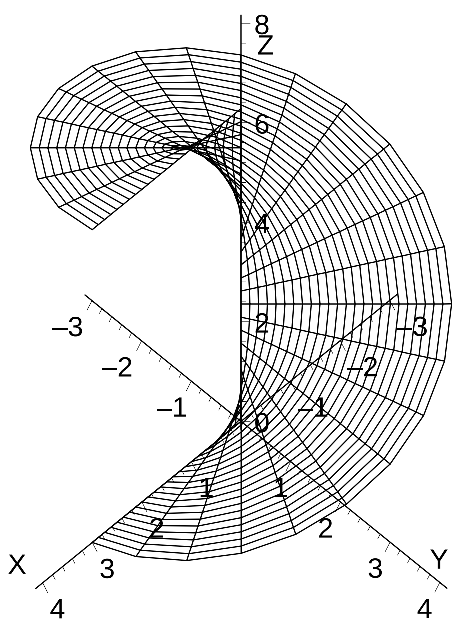


FIGURE 9. Parameter curves on a helicoid

1.3. Tangent Planes and Normal Vectors.

1.3.1. *Tangent Planes.* There are more curves on S than those parameter curves - and we will investigate the geometric features of S mainly by investigating the curves "living on" S. Let again $\mathbf{r} : D \to \mathbf{R}^3$ denote a coordinate patch for S. Obviously, a plane curve consisting of points in D is "translated" via $\mathbf{r}$ to a curve with all points on S. More precisely:

DEFINITION 3.9. A space curve C with parametrization $\mathbf{x} : I \to \mathbf{R}^3$ is called a *smooth* curve *on* $\mathbf{r}(D) \subset S$ if and only if there is a smooth parametrization $(u(t), v(t))$, $t \in I$ of a *plane* curve in D such that

$$\mathbf{x}(t) = \mathbf{r}(u(t), v(t)). \tag{3.2}$$

Intuitively, the plane curve $(u(t), v(t))$ arises as the image of C when the curved coordinate system on S is straightened. In fact, every smooth space curve $\mathbf{x} : I \to \mathbf{R}^3$ with $\mathbf{x}(t) \in S$ for all $t \in I$ is a smooth curve in the sense above, but it is not quite elementary to show that.

Let us calculate the *velocity vector* corresponding to the curve C with parametrization (3.2) at the parameter t_0, i.e., at the point $P_0 \in S$ with $\overrightarrow{OP_0} = \mathbf{x}(t_0)$. Let $(u_0, v_0) = (u(t_0), v(t_0))$. Using the chain rule (cf. (3.1)), we obtain

$$\mathbf{x}'(t_0) = u'(t_0)\mathbf{r}_u(u_0, v_0) + v'(t_0)\mathbf{r}_v(u_0, v_0). \tag{3.3}$$

The remarkable fact is, that – for *every* smooth curve on S through P_0 – the velocity vector $\mathbf{x}'(t_0)$ is contained in $sp(\mathbf{r}_u(u_0, v_0), \mathbf{r}_v(u_0, v_0))$, the linear subspace of $\mathbf{R}^3$ consisting of all linear combinations of the two partial derivative vectors. Intuitively, this means that the tangent vectors to curves through a point P_0 are all contained in the plane spanned by the tangent vectors to the two parameter curves. This plane is the *best approximating plane* to the surface S close to the point P_0.

DEFINITION 3.10. Let S denote a regular surface and $P_0 \in S$.

(1) The *linear tangent plane* $T_{P_0}S$ to S at P_0 consists of all velocity vectors to smooth curves on S through P_0. Given a coordinate patch $\mathbf{r} : D \to \mathbf{R}^3$ for S with $\mathbf{r}(u_0, v_0) = \overrightarrow{OP_0}$, it has a parametrization

$$T_{P_0}S = \{s\mathbf{r}_u(u_0, v_0) + t\mathbf{r}_v(u_0, v_0),\ s, t \in \mathbf{R}\}. \tag{3.4}$$

(2) The *affine tangent plane* π_{P_0S} to S at P_0 consists of all points $Q \in \mathbf{E}^3$ with $\overrightarrow{P_0Q} \in T_{P_0}S$. It has a parametrization

(3.5)
$$\pi_{P_0}S = \{Q \in \mathbf{E}^3 | \overrightarrow{OQ} = \overrightarrow{OP_0} + s\mathbf{r}_u(u_0, v_0) + t\mathbf{r}_v(u_0, v_0),\ s, t \in \mathbf{R}\}.$$

REMARK 3.11.
The affine tangent plane consists of the tangent vectors to S *attached* to the point P_0.

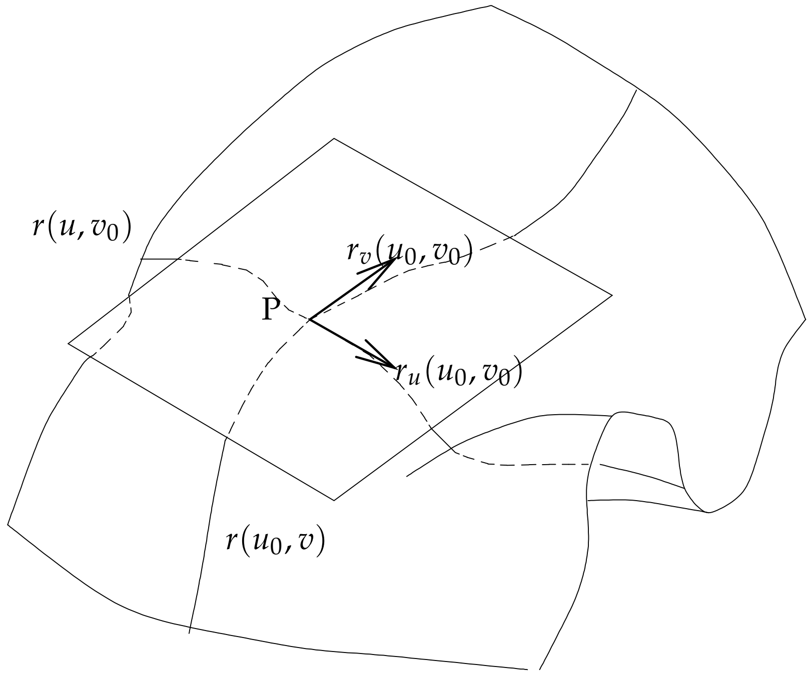


FIGURE 10. The tangent plane

EXAMPLE 3.12.

(1) Let S denote the cylinder from Ex. 3.6.1 with parametrization $\mathbf{r}(u, v) = [\rho \cos v, \rho \sin v, u]$. The partial derivatives were calculated in Ex. 3.8 as

$$\mathbf{r}_u(u_0, v_0) = [0, 0, 1], \quad \mathbf{r}_v(u_0, v_0) = [-\rho \sin v_0, \rho \cos v_0, 0].$$

These vectors are in fact the tangent vectors to the parameter curves we identified in Ex. 3.8. The tangent plane at the point $P_0 \in S$ with $\mathbf{r}(u_0, v_0) = \overrightarrow{OP_0}$ is thus

$$T_{P_0}S = \{s[0, 0, 1] + t[-\rho \sin v_0, \rho \cos v_0, 0], \quad s, t \in \mathbf{R}\}.$$

The affine tangent plane at P_0 is

$$\pi_{P_0}S = \{Q \in \mathbf{E}^3 | \overrightarrow{OQ} = [\rho \cos v_0, \rho \sin v_0, u_0] + s[0, 0, 1] + t[-\rho \sin v_0, \rho \cos v_0, 0], \ s, t \in \mathbf{R}\}.$$

(2) Let S denote the surface obtained as the graph of a smooth function $f : D \to \mathbf{R}$ with parametrization $\mathbf{r}(u, v) = [u, v, f(u, v)]$ as in Ex. 3.6.3. Let $P_0 \in S$ be given by $\overrightarrow{OP_0} = \mathbf{r}(u_0, v_0) = [u_0, v_0, f(u_0, v_0)]$. Since

$$\mathbf{r}_u(u_0, v_0) = [1, 0, f_u(u_0, v_0)] \text{ and } \mathbf{r}_v(u_0, v_0) = [0, 1, f_v(u_0, v_0)],$$

the tangent space T_{P_0} is the two-dimensional subspace

$$sp\{[1, 0, f_u(u_0, v_0)], [0, 1, f_v(u_0, v_0)]\}.$$

1.3.2. *Normal vectors.* Let again S be a regular surface and $P_0 \in S$ a point on S. The tangent plane $T_{P_0}S$ to S at P_0 consists of all tangents at P_0 to smooth curves on S through P_0. Being a plane in $\mathbf{R}^3$ it has a uniquely determined *normal line* consisting of vectors that are perpendicular to the tangent plane:

DEFINITION 3.13. A vector $\mathbf{n} \in \mathbf{R}^3$ is called a *normal vector* to S at P_0 if $\mathbf{n}$ is *perpendicular* to all tangent vectors $\mathbf{v} \in T_{P_0}S$.

Given a coordinate patch $\mathbf{r} : D \to \mathbf{R}^3$ for the surface S with $\mathbf{r}(u_0, v_0) = \overrightarrow{OP}$, it is easy to calculate normal vectors. Since $T_{P_0}S = sp(\mathbf{r}_u(u_0, v_0), \mathbf{r}_v(u_0, v_0))$, we get a normal vector by the cross product of the two spanning vectors, i.e., $\mathbf{n} = \mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0))$. Since we are only interested in the normal direction, it is customary to replace that vector by a *unit* vector with the same direction:

DEFINITION 3.14. Let $\mathbf{r} : D \to \mathbf{R}^3$ denote a coordinate patch for the surface S with $\mathbf{r}(u_0, v_0) = \overrightarrow{OP}$, $P \in S$. The vector

$$\boldsymbol{\nu}(P) = \boldsymbol{\nu}(u_0, v_0) = \frac{\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)}{|\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)|} \in \mathbf{R}^3$$

is a *unit normal vector* to the surface S at the point P.

Unit normal vectors are unique up to sign. The other unit normal vector to S at P is just the vector $-\boldsymbol{\nu}(P)$.

As usual, a normal vector can be used to describe the tangent plane by an equation: Let $\mathbf{n} \in \mathbf{R}^3$ be any normal vector to the surface S at P_0; $[x_0, y_0, z_0]$. Then a point $Q : [x, y, z] \in \mathbf{R}^3$ is contained in the affine tangent plane $\pi_{P_0}S$ if and only if it satisfies the equation

$$0 = [x - x_0, y - y_0, z - z_0] \cdot \mathbf{n}. \tag{3.6}$$

In particular, let S be the surface determined by the parametrization $\mathbf{r}$ and let $P_0 \in S$ be given by $\overrightarrow{OP_0} = \mathbf{r}(u_0, v_0) = [x(u_0, v_0), y(u_0, v_0), z(u_0, v_0)]$. The vector $\mathbf{n}(u_0, v_0) = \mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)$ determines a normal vector at P_0. Hence, we obtain

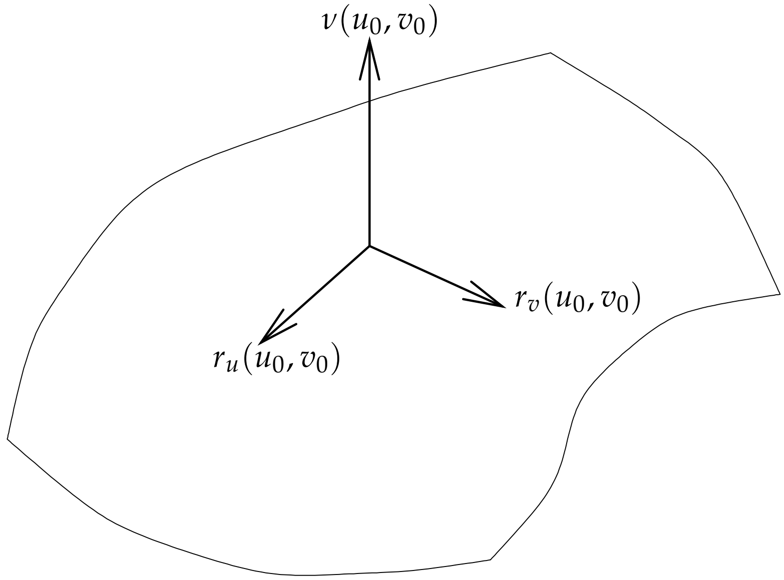


FIGURE 11. Surface with normal

the following forms for an equation for the affine tangent plane at P_0:

$$0 = [x - x(u_0,v_0), y - y(u_0,v_0), z - z(u_0,v_0)] \cdot (\mathbf{r}_u(u_0,v_0) \times \mathbf{r}_v(u_0,v_0)) \tag{3.7}$$

$$= [x - x(u_0,v_0), y - y(u_0,v_0), z - z(u_0,v_0), \mathbf{r}_u(u_0,v_0), \mathbf{r}_v(u_0,v_0)] \tag{3.8}$$

$$= \begin{vmatrix} x - x(u_0,v_0) & y - y(u_0,v_0) & z - z(u_0,v_0) \\ x_u(u_0,v_0) & y_u(u_0,v_0) & z_u(u_0,v_0) \\ x_v(u_0,v_0) & y_v(u_0,v_0) & z_v(u_0,v_0) \end{vmatrix}. \tag{3.9}$$

This last version uses the space product as a combination of dot product and cross product, cf. Sect. 1.2.

In practical terms, we obtain the following recipe for determining an equation for the affine tangent plane of surface S at point P_0: Given a coordinate patch $\mathbf{r}$ for S such that

- $\overrightarrow{OP_0} = \mathbf{r}(u_0, v_0) = [x_0, y_0, z_0]$, a point on S and on the tangent plane $\pi_{p_0} S$.
- Determine $[a, b, c] = \mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)$ as a normal vector to the tangent plane $\pi_{p_0} S$.

Hence, the formula given in (1.22) explains that the following is an equation for the affine tangent plane $\pi_{p_0} S$ at P_0:

$$a(x - x_0) + b(y - y_0) + c(z - z_0)00.$$

EXAMPLE 3.15.
Let S again denote the cylinder with parametrization $\mathbf{r}(u, v) = [\rho \cos v, \rho \sin v, u]$ from

Ex. 3.6 with partial derivatives $\mathbf{r}_u(u_0, v_0) = [0, 0, 1]$, $\mathbf{r}_v(u_0, v_0) = [-\rho \sin v, \rho \cos v, 0]$, cf. Ex. 3.12. We calculate the normal vector $\nu(P_0)$ at the point P_0 with $\mathbf{r}(u_0, v_0) = \overrightarrow{OP_0}$ as

$$\nu(P_0) = \frac{1}{\rho}[0, 0, 1] \times [-\rho \sin v_0, \rho \cos v_0, 0] =$$

$$= \frac{1}{\rho}[-\rho \cos v_0, -\rho \sin v_0, 0] = [-\cos v_0, -\sin v_0, 0].$$

This normal vector points horizontally inwards from the point in the direction of the Z-axis.

A point $Q : [x, y, z] \in \mathbf{R}^3$ is contained in the affine tangent plane $\pi_{P_0} S$ if and only if it satisfies the equation

$$0 = [x - \rho \cos v_0, y - \rho \sin v_0, z - u_0] \cdot [-\cos v_0, -\sin v_0, 0],$$

i.e., $(\cos v_0)x + (\sin v_0)y = \rho$.

2. Measurement in curved coordinates: the 1. fundamental form

2.1. Arc length in curved coordinates: the "curved Pythagoras". Our next aim is to measure lengths, angles and areas. Of course, we know how to do that in 3-dimensional space. But now we are only allowed to do measurements *on the given surface* – using one or several (chosen) parametrization(s). Why that restriction? Well, the shortest path from one point on the surface of the earth to another is *not* given by a straight line segment (through the earth's interior) but by a curve – with curvature. (Which?) How would you draw such a curve using latitude and longitude as coordinates? How do we measure lengths, angles and areas in given *curved coordinates*? To find out, let us first look at the formulas known for space curves and translate the results into our setting using a given parametrization for the surface under investigation.

For a space curve with parametrization $\mathbf{x}(t)$, length is calculated by integrating its speed $|\mathbf{x}'(t)|$, cf. Def. 2.20. Let now $\mathbf{r} : D \to \mathbf{R}^3$ denote a coordinate patch for a surface S. If we look at a curve in that coordinate patch given as $\mathbf{x}(t) = \mathbf{r}(u(t), v(t))$, then its derivative (3.1) is

$$\mathbf{x}'(t) = u'(t)\mathbf{r}_u(u(t), v(t)) + v'(t)\mathbf{r}_v(u(t), v(t)),$$

expanded in the basis $\{\mathbf{r}_u, \mathbf{r}_v\}$ of the tangent space at the relevant point. But these two (moving) basis vectors are, in general, *neither unit vectors nor orthogonal to each other*. So, the length of a vector

$$\mathbf{v} = a\mathbf{r}_u(u_0, v_0) + b\mathbf{r}_u(u_0, v_0)$$

expanded in this basis cannot just be calculated by the usual plane Pythagoras theorem. Instead we need what might be called a "curved Pythagoras theorem":

DEFINITION 3.16. Let $\mathbf{r} : D \to \mathbf{R}^3$ denote a coordinate patch for the surface S. We define three functions $E, F, G : D \to \mathbf{R}$ given by

$$\begin{aligned} E(u,v) &= \mathbf{r}_u(u,v) \cdot \mathbf{r}_u(u,v) \\ F(u,v) &= \mathbf{r}_u(u,v) \cdot \mathbf{r}_v(u,v) \\ G(u,v) &= \mathbf{r}_v(u,v) \cdot \mathbf{r}_v(u,v). \end{aligned} \tag{3.10}$$

REMARK 3.17.

(1) How can one interpret the functions E, F and G? Let us look at the parameter curves through the point $P \in S$ with $\overrightarrow{OP} = \mathbf{r}(u_0, v_0)$. Then,
$E(u_0, v_0) = |\mathbf{r}_u(u_0, v_0)|^2$ is the square of the *speed* along the 1. parameter curve at P, and likewise, $G(u_0, v_0)$ is the square of the speed along the 2. parameter curve at P. In particular,

$$E(u_0, v_0) > 0,\ G(u_0, v_0) > 0. \tag{3.11}$$

Moreover, the two parameter curves have an *orthogonal intersection* at P if and only if $F(u_0, v_0) = \mathbf{r}_u(u_0, v_0) \cdot \mathbf{r}_v(u_0, v_0) = 0$.

(2) For the parametrization of the plane given as

$$\mathbf{r} : \mathbf{R}^2 \to \mathbf{R}^3,\ \mathbf{r}(u,v) = (u, v, 0),$$

we calculate at any point $(u_0, v_0) \in \mathbf{R}^2$:

$$\mathbf{r}_u(u_0, v_0) = [1,0,0],\ \mathbf{r}_v(u_0, v_0) = [0,1,0], \text{ and thus}$$

$$\begin{aligned} E(u_0, v_0) = [1,0,0] \cdot [1,0,0] &= 1, \\ F(u_0, v_0) = [1,0,0] \cdot [0,1,0] &= 0, \\ G(u_0, v_0) = [0,1,0] \cdot [0,1,0] &= 1. \end{aligned}$$

LEMMA 3.18. *Let $P \in S$ denote a point on the surface S with $\overrightarrow{OP} = \mathbf{r}(u_0, v_0)$. The length of the tangent vector $\mathbf{w} = a\mathbf{r}_u(u_0, v_0) + b\mathbf{r}_v(u_0, v_0) \in T_pS$ is given by*

$$|\mathbf{w}|^2 = a^2 E(u_0, v_0) + 2ab F(u_0, v_0) + b^2 G(u_0, v_0).$$

PROOF:
The proof is an easy calculation using the dot product and its properties (cf. Sect. 1.2.6). We shall write $\mathbf{r}_u$ and $\mathbf{r}_v$ to abbreviate $\mathbf{r}_u(u_0, v_0)$, resp. $\mathbf{r}_v(u_0, v_0)$:

$$\begin{aligned} |\mathbf{w}|^2 &= \mathbf{w} \cdot \mathbf{w} \\ &= (a\mathbf{r}_u + b\mathbf{r}_v) \cdot (a\mathbf{r}_u + b\mathbf{r}_v) \\ &= a^2(\mathbf{r}_u \cdot \mathbf{r}_u) + 2ab\mathbf{r}_u \cdot \mathbf{r}_v + b^2(\mathbf{r}_v \cdot \mathbf{r}_v) \\ &= a^2E + 2abF + b^2G. \qquad (3.12) \end{aligned}$$

□

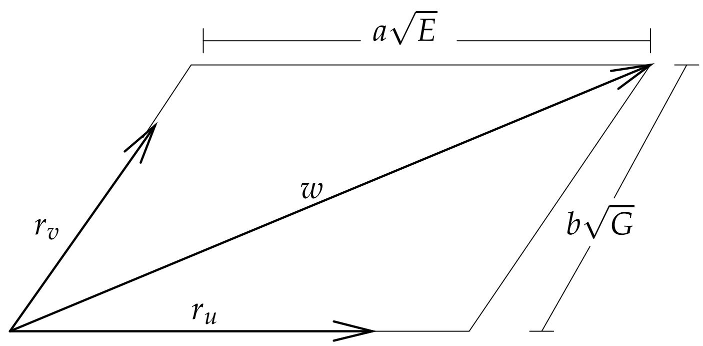


FIGURE 12. The "curved Pythagoras theorem"

REMARK 3.19.

(1) For the parametrization of the plane from Rem. 3.17, the length of $\mathbf{w} = a\mathbf{r}_u + b\mathbf{r}_v$ is given by
$$|\mathbf{w}|^2 = a^2E + 2abF + b^2G = a^2 + b^2.$$
So, Lemma 3.18 has *Pythagoras theorem* as a special case.

(2) Here is how to interpret Lemma 3.18 in general coordinates: In calculating lengths of tangent vectors you have to compensate for the fact that, in general, parameter curves are not perpendicular to each other nor given in arc length parametrization. So, the coefficients E, F, and G in Lemma (3.18) can be interpreted as correction factors to the classical Pythagoras theorem: E and G correct for the speed along parameter curves, F corrects for the angle between those.

To calculate the length of a curve on a surface, remember that the arc length function $s(t)$ is defined as the integral over the speed (i.e., length of velocity vector) along that curve. This yields the following formula:

PROPOSITION 3.20. *Let $\mathbf{r} : D \to \mathbf{R}^3,\ D \subset \mathbf{R}^2$ open, denote a coordinate patch for the surface S; let furthermore $\mathbf{x}(t) = \mathbf{r}(u(t), v(t)),\ a \le t \le b$ denote a parametrization for a curve C on S. The derivative $s'(t) = \frac{ds}{dt}$ of the arc length function $s(t),\ a \le t \le b$ (cf. 2.20) satisfies:*

(3.13)

$$(s'(t))^2 = u'(t)^2 E(u(t), v(t)) + 2u'(t)v'(t)F(u(t), v(t)) + (v'(t))^2 G(u(t), v(t)).$$

and thus the length of the segment of the curve C between the points corresponding to the parameters t_0 to t_1 is given (in short form) by:

$$s(t_1) = \int_{t_0}^{t_1} s'(t)\, dt$$

$$= \int_{t_0}^{t_1} \sqrt{(u')^2 E + 2u'v'F + (v')^2 G}\, dt. \tag{3.14}$$

REMARK 3.21.

(1) In the literature, you will often find the following short form for the length formula (3.13) in curved coordinates:

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2.$$

It is often called the first *fundamental form*, and E, F and G are called the *coefficients* of the 1. fundamental form. The first fundamental form may be considered as a quadratic polynomial in two variables defined on tangent vectors $\mathbf{t} = a\mathbf{r}_u + b\mathbf{r}_v$, to wit,

$$I(\mathbf{t}) = Ea^2 + 2Fab + Gb^2. \tag{3.15}$$

In general, the first fundamental form is *not* constant but varies from point to point (as in Ex. 3.22.2 below); moreover, the *expression* for it as function of the parameters u and v depends on the parametrization.

(2) As for the general arc length formula, it is often difficult or impossible to use (3.14) to calculate the length of a concrete surface curve exactly. The main value of Prop. 3.20 is that it allows us to *reason* about arc length, and later on to use this in actual calculations of surface curvatures.

EXAMPLE 3.22.

(1) Let $\mathbf{r}(u,v) = [\rho\cos v, \rho\sin v, u]$ denote the parametrization for a cylinder from Ex. 3.6.1. We found earlier (cf. Ex. 3.12):

$$\mathbf{r}_u(u,v) = [0,0,1] \text{ and } \mathbf{r}_v(u,v) = [-\rho\sin v, \rho\cos v, 0],$$

and can thus calculate:

$$\begin{aligned} E(u,v) &= [0,0,1]\cdot[0,0,1] = 1; \\ F(u,v) &= [0,0,1]\cdot[-\rho\sin v, \rho\cos v, 0] = 0; \\ G(u,v) &= [-\rho\sin v, \rho\cos v, 0]\cdot[-\rho\sin v, \rho\cos v, 0] = \rho^2. \end{aligned}$$

Remark, that the parameter curves on the cylinder intersect each other at a right angle everywhere ($F \equiv 0$). By Prop. 3.20, the length of any curve $\mathbf{x}(t) = \mathbf{r}(u(t), v(t))$ in cylindrical coordinates $(u(t), v(t))$ from parameter t_0 to parameter t_1 is:

$$\int_{t_0}^{t_1} \sqrt{u'(t)^2 + \rho^2 v'(t)^2}dt.$$

(2) Let $\mathbf{r}(u,v) = [\rho\sin u\cos v, \rho\sin u\sin v, \rho\cos u]$ denote the parametrization for a sphere with radius ρ from Ex. 3.6.2. The partial derivatives of $\mathbf{r}$ are calculated as:

$$\begin{aligned} \mathbf{r}_u(u,v) &= \rho[\cos u\cos v, \cos u\sin v, -\sin u]; \\ \mathbf{r}_v(u,v) &= \rho[-\sin u\sin v, \sin u\cos v, 0], \text{ and thus} \end{aligned}$$

$$\begin{aligned} E(u,v) &= \rho^2[\cos u\cos v, \cos u\sin v, -\sin u]\cdot[\cos u\cos v, \cos u\sin v, -\sin u] \\ &= \rho^2((\cos u)^2(\cos v)^2 + (\cos u)^2(\sin v)^2 + (\sin u)^2) = \\ &= \rho^2((\cos u)^2((\cos v)^2 + (\sin v)^2) + (\sin u)^2) = \rho^2((\cos u)^2 + (\sin u)^2) = \rho^2; \\ F(u,v) &= \rho^2[\cos u\cos v, \cos u\sin v, -\sin u]\cdot[-\sin u\sin v, \sin u\cos v, 0] \\ &= \rho^2(-\cos u\cos v\sin u\sin v + \cos u\sin v\sin u\cos v) = 0; \\ G(u,v) &= \rho^2[-\sin u\sin v, \sin u\cos v, 0]\cdot[-\sin u\sin v, \sin u\cos v, 0] \\ &= \rho^2((\sin u)^2(\sin v)^2 + (\sin u)^2(\cos v)^2) \\ &= \rho^2(\sin u)^2((\sin v)^2 + (\cos v)^2) = \rho^2(\sin u)^2. \end{aligned}$$

Since $F \equiv 0$ everywhere, we have verified that the parameter curves (circles of a given latitude, resp. longitude) intersect each other at a right angle. The length of any curve $\mathbf{x}(t) = \mathbf{r}(u(t), v(t))$ in spherical coordinates $(u(t), v(t))$ from parameter t_0 to parameter t_1 is:

$$\int_{t_0}^{t_1} \rho\sqrt{u'(t)^2 + v'(t)^2(\sin u(t))^2}\, dt.$$

(3) The graph of a smooth function $f : D \to \mathbf{R},\ D \subset \mathbf{R}^2$ open, has the parametrization $\mathbf{r} : D \to \mathbf{R}^3,\ \mathbf{r}(u,v) = [u, v, f(u,v)]$, cf. Ex. 3.6.3. In this case,

$$\mathbf{r}_u(u,v) = [1, 0, f_u(u,v)] \text{ and } \mathbf{r}_v(u,v) = [0, 1, f_v(u,v)],$$

and hence

$$\begin{aligned} E(u,v) &= [1,0,f_u(u,v)] \cdot [1,0,f_u(u,v)] = 1 + (f_u(u,v))^2; \\ F(u,v) &= [1,0,f_u(u,v)] \cdot [0,1,f_v(u,v)] = f_u(u,v) f_v(u,v); \\ G(u,v) &= [0,1,f_v(u,v)] \cdot [0,1,f_v(u,v)] = 1 + (f_v(u,v))^2. \end{aligned}$$

2.2. Angle and Area Measurement. In fact, the first fundamental form gives us not only formulas for lengths but also for angles and areas. This should not be too surprising: In fact, the first fundamental form expresses the dot product between tangent vectors. So, we have to adapt formulas expressing angles by dot products (cf. 1.6) and the area of parallelograms by cross products (cf. 1.2.5 and 1.12). This is what we do now: Let S be (part of) a surface given by a parametrization $\mathbf{r} : D \to \mathbf{R}^3$, $D \subset \mathbf{R}^2$ open.

2.2.1. *Angles.* We look at two surface curves C_1 and C_2 given by parametrizations

$$\mathbf{x}_1(t) = \mathbf{r}(u_1(t), v_1(t)) \text{ and } \mathbf{x}_2(t) = \mathbf{r}(u_2(t), v_2(t)).$$

Suppose furthermore that $(u_1(t_1), v_1(t_1)) = (u_2(t_2), v_2(t_2))$ for certain values t_1 and t_2. This means, that the two curves intersect in a common point P. To find out the size of the angle of intersection at P (between the tangent vectors of the two curves), we apply (1.6). The following formula is in short form, i.e., one has to plug in the relevant parameters t_1, t_2, $(u_1(t_1), v_1(t_1))$, resp. $(u_2(t_2), v_2(t_2))$:

LEMMA 3.23. *The angle α between the curves C_1 and C_2 satisfies:*

$$\cos\alpha = \frac{u_1' u_2' E + (u_1' v_2' + u_2' v_1') F + v_1' v_2' G}{\sqrt{((u_1')^2 E + 2u_1' v_1' F + (v_1')^2 G)((u_2')^2 E + 2u_2' v_2' F + (v_2')^2 G)}}.$$

PROOF:
Apply (1.6) to

$$\mathbf{x} = \mathbf{x}_1'(t_1) = u_1' \mathbf{r}_u + v_1' \mathbf{r}_v \text{ and } \mathbf{y} = \mathbf{x}_2'(t_2) = u_2' \mathbf{r}_u + v_2' \mathbf{r}_v.$$

□

COROLLARY 3.24. *The angle α between the parameter curves at a point P with $\overrightarrow{OP} = \mathbf{r}(u_0, v_0)$ satisfies:*

$$\cos\alpha = \frac{F(u_0, v_0)}{\sqrt{EG(u_0, v_0)}}.$$

EXAMPLE 3.25.
The formulas for the coefficients E, F and G in the first fundamental form of a graph surface S (cf. Ex. 3.22.3) show: The angle α between the parameter curves

$$\mathbf{x}_1(t) = [u, v_0, f(u, v_0)] \text{ and } \mathbf{x}_2(t) = [u_0, v, f(u_0, v)]$$

at the point $P \in S$ with $\overrightarrow{OP} = \mathbf{r}(u_0, v_0)$ satisfies

$$\cos\alpha = \frac{f_u(u_0, v_0) f_v(u_0, v_0)}{\sqrt{(1 + (f_u(u_0, v_0))^2)(1 + (f_v(u_0, v_0))^2)}}.$$

2.2.2. *Area.* We investigate the following situation: Given a region R on the surface S contained in the image of the parametrization $\mathbf{r} : D \to \mathbf{R}^2$. How can one express the area $a(R)$ in terms of the parametrization? First of all, the region R corresponds to a subset of the parameter set D in the plane. More precisely, the set of points $(u_0, v_0) \in D$ with the property that $\mathbf{r}(u_0, v_0)$ points to a point in R is by definition just the set

$$\mathbf{r}^{-1}(R) = \{(u, v) \in D | \mathbf{r}(u, v) \in R\}.$$

Remember that the area of a region R' in the Euclidean plane $\mathbf{E}^2$ can be expressed by a double integral, i.e.,

$$a(R') = \int\int_{R'} 1\, du dv.$$

Our aim is to express the area of $R \subset S$ as a *plane integral* over the subset $\mathbf{r}^{-1}(R) \subset \mathbf{R}^2$. But in general, R and $\mathbf{r}^{-1}(R)$ have very different areas, so we will have to integrate a non-constant function (in two variables) over R'. Let us first give the result (in form of a definition) before motivating it:

DEFINITION 3.26. Let $\mathbf{r} : D \to \mathbf{R}^3$ denote a smooth parametrization for the surface S, and let $R \subset S$ denote a subset whose boundary is a piecewise smooth curve $C \subset S$. The *area* $a(R)$ of R is given as

$$a(R) = \int\int_{\mathbf{r}^{-1}(R)} |\mathbf{r}_u \times \mathbf{r}_v| du dv = \int\int_{\mathbf{r}^{-1}(R)} \sqrt{EG - F^2} du dv. \tag{3.16}$$

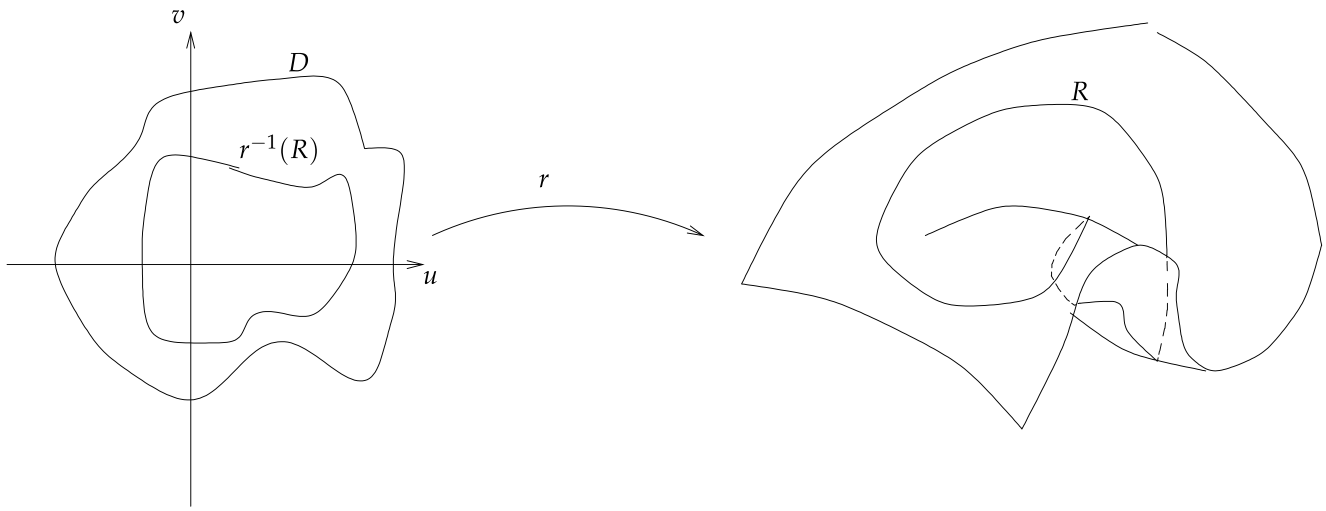


FIGURE 13. Corresponding regions in the parameter plane and on the surface

REMARK 3.27.
Let us first indicate why the two functions to be integrated coincide: Using 1.12 with $\mathbf{x} = \mathbf{r}_u$ and $\mathbf{y} = \mathbf{r}_v$ and the definitions of E, F and G (cf. 3.16), we obtain:

$$EG = |\mathbf{r}_u|^2|\mathbf{r}_v|^2 = |\mathbf{r}_u \times \mathbf{r}_v|^2 + (\mathbf{r}_u \cdot \mathbf{r}_v)^2 = |\mathbf{r}_u \times \mathbf{r}_v|^2 + F^2.$$

An important consequence of this and the definition of a parametrization (Def. 3.4.2) is:

$$(EG - F^2)(u, v) > 0 \text{ for all } (u, v) \in D.$$

EXAMPLE 3.28.
Let us calculate the area of the region R on a sphere with radius ρ with spherical coordinates

$$u_0 \leq u \leq u_1; \quad v_0 \leq v \leq v_1.$$

We use the parametrization

$$\mathbf{r} :]0, \pi[\times]0, 2\pi[\to \mathbf{R}^3, \ \mathbf{r}(u, v) = [\rho \sin u \cos v, \rho \sin u \sin v, \rho \cos u]$$

for the sphere from Ex. 3.6.2 with the following coefficients for the first fundamental form (cf. 3.22):

$$E = \rho^2; \quad F = 0; \quad G = \rho^2(\sin u)^2,$$

and hence, $(EG - F^2) = \rho^4(\sin u)^2$. Furthermore, the plane set

$$\mathbf{r}^{-1}(R) = \{(u, v) \in \mathbf{R}^2 | u_0 \leq u \leq u_1; \ v_0 \leq v \leq v_1\}$$

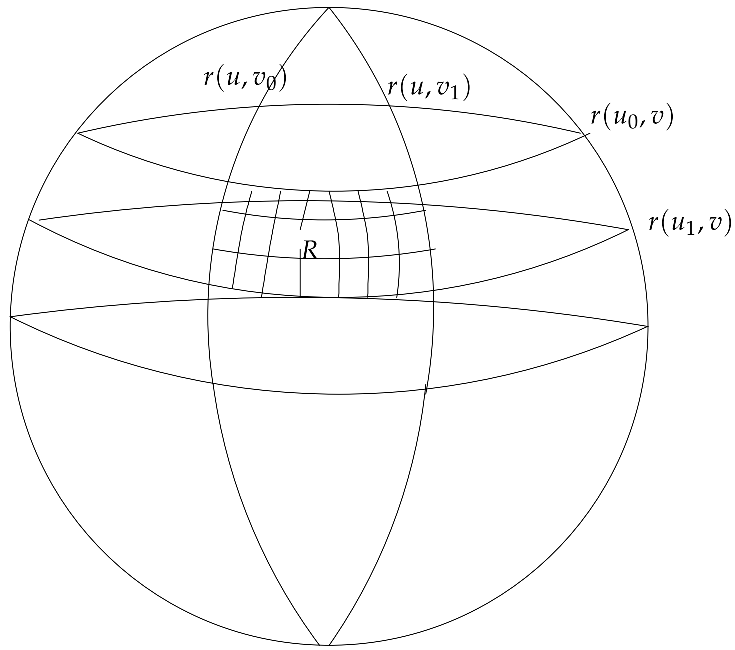


FIGURE 14. A spherical region

is just a rectangle. Hence,

$$a(R) = \int\int_{\mathbf{r}^{-1}(R)} \rho^2 |\sin u| \, dudv = \rho^2 \int_{u_0}^{u_1} (\int_{v_0}^{v_1} |\sin u| \, dv) du =$$

$$\rho^2 \int_{u_0}^{u_1} (v_1 - v_0) \sin u du = \rho^2 (v_1 - v_0)(\cos u_0 - \cos u_1).$$

This formula allows us to calculate the area of the complete sphere $S(\rho)$, corresponding to $u_0 = 0$, $u_1 = \pi$, $v_0 = 0$, $v_1 = 2\pi$. Its area is:

$$a(S(\rho)) = \rho^2 (2\pi - 0)(1 - (-1)) = 4\pi\rho^2,$$

as you probably know in advance.

Here is a *motivation* for the area formula (3.16): We look at a partition of the parameter region $\mathbf{r}^{-1}(R)$ into small squares of side length h. The images of these squares partition the region R within our surface S. Every such image is bounded by four surface curves, that are segments of the 1. resp. the second parameter curve, cf. Fig. 15.

Let us now look at one specific square Q in the plane with vertices $(u_0, v_0), (u_0 + h, v_0), (u_0, v_0 + h)$ and $(u_0 + h, v_0 + h)$ and area $a(Q) = h^2$. The area of the image $\mathbf{r}(Q)$ is more difficult to get hold on. It is a small region $\mathbf{r}(Q)$ with corner points P_0, P_1, P_2 and P_3, the images of the four vertices above. Let us now approximate this small region $\mathbf{r}(Q)$ by a parallelogram. The best approximating parallelogram Q' with vertex P_0 is spanned by the tangent vectors to the parameter curves at the point P_0 with

$\overrightarrow{OP_0} = \mathbf{r}(u_0, v_0)$. The definition of the partial derivatives from Calculus tells us that – for small $h > 0$ – the point P_1' with $\overrightarrow{OP_1'} = \mathbf{r}(u_0, v_0) + h\mathbf{r}_u(u_0, v_0)$ is a good approximation to P_1; likewise P_2 is approximated by P_2' with $\overrightarrow{OP_2'} = \mathbf{r}(u_0, v_0) + h\mathbf{r}_v(u_0, v_0)$. Since

$$\overrightarrow{P_0P_1'} = h\mathbf{r}_u(u_0, v_0) \text{ and } \overrightarrow{P_0P_2'} = h\mathbf{r}_v(u_0, v_0),$$

the area of the parallelogram Q' is given by (cf. Sct. 1.2.5):

$$a(Q') = |h\mathbf{r}_u(u_0, v_0) \times h\mathbf{r}_v(u_0, v_0)| = h^2|\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)|,$$

and thus the scaling (or magnification) factor is

$$\frac{a(Q')}{a(Q)} = \frac{h^2|\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)|}{h^2} = |\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)|.$$

This scaling factor is then an *approximate scaling* factor comparing the areas of the square Q and its image $\mathbf{r}(Q)$. This may be expressed as

$$a(\mathbf{r}(Q)) \approx h^2|\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)| = \int\int_Q |\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0)| du dv$$

– as the integral of a constant function.

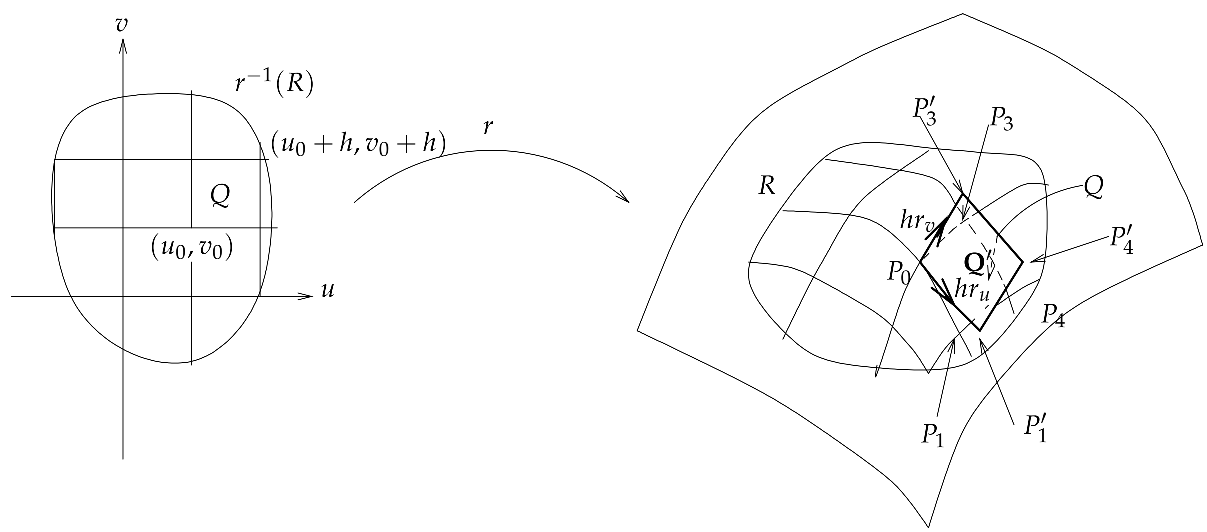


FIGURE 15. Partitions and approximations

Finally, we have to sum over *all* the squares in the partition and to make the partition finer and finer, i.e., we let h tend to 0. The finite sum is then replaced by a plane integral, and the function to be integrated is the function $|\mathbf{r}_u \times \mathbf{r}_v|$ from Def. 3.26. The value of that function at a parameter point has thus the interpretation as a local *area scaling factor* – which, in general, varies from point to point.

REMARK 3.29.

(1) The explanation above is not rigourous, and, in fact, the definition of areas is one of the more subtle points in geometry. For a more solid foundation, one would have to find out that the limit proces above in fact converges, that it converges to the integral above, and this includes a more thorough look at the boundary of the region. Moreover, one needs that the area formula is independent of the chosen parametrization. This can be shown using a transformation formula for plane integrals from Advanced Calculus.
(2) Here is an alternative approach leading to formula (3.16): We construct a "normal thickening" $T_h(R)$ of the region R, a slice around R in normal direction of "width" h. This thickening has a *volume* $v(T_h(R))$, related to the area of R by $a(R) \approx \frac{v(T_h(R))}{h}$. More precisely, $a(R) = \lim_{h\to\infty} \frac{v(T_h(R))}{h}$. The calculation of volumes for regions in $\mathbf{E}^3$ can be performed using the space product (cf. Sect. 1.2.4).

3. Normal sections and normal curvature

The main aim with this chapter is to motivate, define and calculate various notions of curvature on a surface. First of all, curvature properties for a surface are more complex than those for a curve. The simple reason is that, from a point P on the surface, you can look in (infinitely) many surface directions, and the curvature might vary between all these directions. It would be painful to have to handle infinitely many curvatures. Nevertheless, this is what we do to start with; in the end, it turns out, that we only need *two* numbers, at very point P, to characterise curvature on the surface at the given point.

3.1. Definitions. Let S be a surface, and let P be a point on S. At P, the surface has a tangent plane T_pS consisting of all velocity vectors of curves *on* S *through* P, cf. Def. 3.10. Choose a *unit normal vector* $\nu(P)$ (cf. Def. 3.14) , and fix a *unit tangent vector* $\mathbf{v} \in T_pS$, i.e., $|\mathbf{v}| = 1$. (This vector then specifies a tangent *direction*). These two vectors span the *normal plane* $\pi_{\mathbf{v}}(P)$ through P (with a parametrization

$$\overrightarrow{OQ} = \overrightarrow{OP} + s\mathbf{v} + t\boldsymbol{\nu}(P), \quad s, t \in \mathbf{R}).$$

The intersection of this normal plane with the surface S is a surface curve called the *normal section* in the direction $\mathbf{v}$:

DEFINITION 3.30. The *normal section* $C_{\mathbf{v}}(P)$ of the surface S at the point $P \in S$ in direction $\mathbf{v} \in T_P S$ is the curve which arises as the intersection of the normal plane $\pi_{\mathbf{v}}(P)$ and the surface S, i.e.,

$$C_{\mathbf{v}}(P) = S \cap \pi_{\mathbf{v}}(P).$$

The *normal curvature* $k_n(P; \mathbf{v})$ is then defined as the (*plane*) curvature of the normal section $C_{\mathbf{v}}(P)$ viewed as a curve in the normal plane $\pi_{\mathbf{v}}(P)$ with orientation given by the basis $\{\mathbf{v}, \boldsymbol{\nu}(P)\}$.

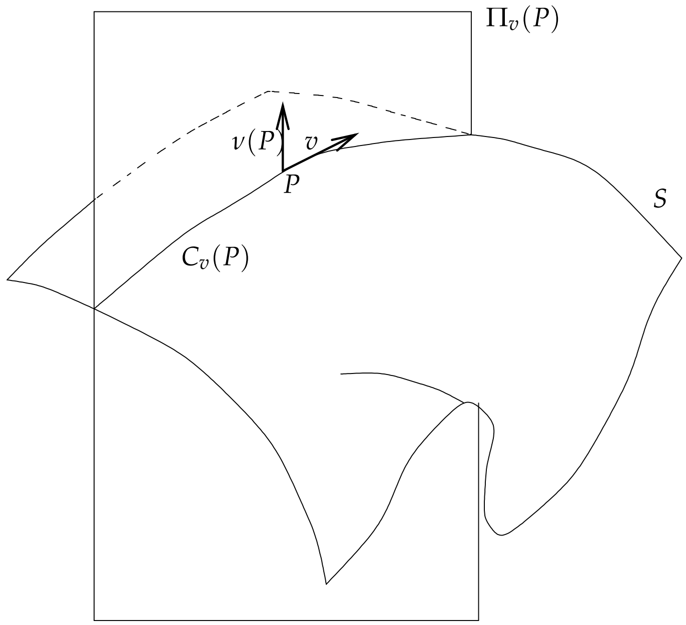


FIGURE 16. Surface, normal plane, and normal section

More precisely, let $\mathbf{r}_{\mathbf{v}} : I \to S$ denote the arc length parametrization (cf. 2.22) of the normal section $C_{\mathbf{v}}(P)$ with $\mathbf{r}_{\mathbf{v}}(0) = \overrightarrow{OP}$ oriented in such a way that $\mathbf{r}'_{\mathbf{v}}(0) = \mathbf{v}$. Moreover,

$$\mathbf{r}''_{\mathbf{v}}(0) = k_n(P; \mathbf{v}) \boldsymbol{\nu}(P). \tag{3.17}$$

Remark, that $\mathbf{r}''_{\mathbf{v}}(0)$ is perpendicular to $\mathbf{r}'_{\mathbf{v}}(0) = \mathbf{v}$ and contained in the plane $sp\{\mathbf{v}, \boldsymbol{\nu}(P)\}$, so (3.17) makes sense. In particular, the normal curvature

$$k_n(P; \mathbf{v}) = \mathbf{r}''_{\mathbf{v}}(0) \cdot \boldsymbol{\nu}(P). \tag{3.18}$$

can be both positive, negative, or zero. For a normal section with a parametrization not necessarily by arc length, the normal curvature takes the form

$$k_n(P;\mathbf{v}) = \frac{\mathbf{r}''_{\mathbf{v}}(0)\cdot \boldsymbol{\nu}(P)}{|\mathbf{r}'_{\mathbf{v}}(0)|^3}. \tag{3.19}$$

3.2. Calculation of the normal curvature in a special case. The normal sections slice the surface close to a point P, and the normal curvatures at P measure the curvatures of the resulting curves. While these infinitely many numbers give a picture of the curvature on the surface near S in *every* tangential direction, it would be painful if one had to calculate every single normal curvature. Is there a *hidden pattern* among these normal curvatures? Is it possible to do fewer (and in particular only finitely many) calculations but still to get the entire picture? A preliminary answer to these questions will first be given in a *special case*; later on, it will turn out, that the answer is similar in the general case:

The special case is that of a surface S obtained as the *graph of a smooth function* $f : D \to \mathbf{R},\ D \subset \mathbf{R}^2$ open and $\mathbf{0} = [0,0] \in D$. Suppose moreover:

$$f(0,0) = f_x(0,0) = f_y(0,0) = f_{xy}(0,0) = 0.$$

The first condition ensures that the surface S contains the origin $O \in \mathbf{E}^3$. The next two conditions entail that the tangent plane T_OS at O is spanned by $[1,0,0]$ and $[0,1,0]$ (cf. Ex. 3.12.2). Hence, the affine tangent plane π_OS coincides with the XY-plane, and $\mathbf{k} = [0,0,1] \in \mathbf{R}^3$ can be chosen as unit normal vector: $\boldsymbol{\nu}(O) = \mathbf{k}$.

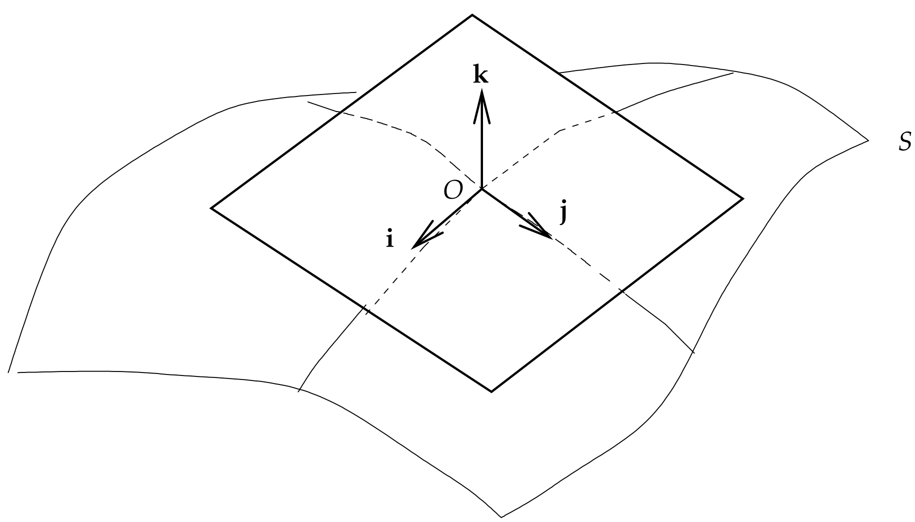


FIGURE 17. A surface with horizontal tangent plane

It is always possible to move any given surface S and any given point $P \in S$ by a rigid motion (i.e., without changing curvature properties) to a new position that has the properties above: A parallel translation moves the point P into O. A first rotation (through O) rotates the normal vector into $\mathbf{k}$ and thus the tangent plane into horizontal position. A second rotation with axis $\mathbf{k}$ can then be performed to meet the last condition $f_{xy}(0,0) = 0$. This can be shown by the methods from Sect. 4.

Let us find the normal sections *in this special case*: A unit tangent vector $\mathbf{v} \in T_O S$ is of the form $\mathbf{v} = [\cos\theta, \sin\theta, 0]$ for some angle $\theta \in [0, 2\pi]$. We will use this angle θ to parameterize the normal sections $C(\theta) = C_{\mathbf{v}}(O)$ and the normal curvatures $k_n(\theta) = k_n(P; \mathbf{v})$. Now, the normal section $C(\theta)$ is the intersection of the plane $sp\{[\cos\theta, \sin\theta, 0], \mathbf{k}]\} = \{[t\cos\theta, t\sin\theta, s] |\ s, t \in \mathbf{R}\}$ with $S = \{[x, y, f(x, y)] |\ (x, y) \in D\}$. Since $x = t\cos\theta$ and $y = t\sin\theta$, we deduct that $z = s = f(x, y) = f(t\cos\theta, t\sin\theta)$ for a point $[x, y, z]$ in the intersection $C(\theta)$. The normal section curve $C(\theta)$ can therefore be parametrized by the vector function

$$\mathbf{r} : I \to \mathbf{R}^3, \quad \mathbf{r}(t) = [t\cos\theta, t\sin\theta, f(t\cos\theta, t\sin\theta)].$$

Let us now do the neccessary calculations – using the chain rule 3.1 – in order to obtain the normal curvature in direction θ:

$$\mathbf{r}'(t) = [\cos\theta, \sin\theta, \cos\theta f_x(t\cos\theta, t\sin\theta) + \sin\theta f_y(t\cos\theta, t\sin\theta)];$$

$$\begin{aligned} \mathbf{r}''(t) &= [0, 0, (\cos\theta)^2 f_{xx}(t\cos\theta, t\sin\theta) + 2\cos\theta\sin\theta f_{xy}(t\cos\theta, t\sin\theta) \\ &\quad +(\sin\theta)^2 f_{yy}(t\cos\theta, t\sin\theta)], \end{aligned}$$

and, in particular,

$$\begin{aligned} \mathbf{r}'(0) &= [\cos\theta, \sin\theta, 0] = \mathbf{v}; \\ \mathbf{r}''(0) &= [0, 0, (\cos\theta)^2 f_{xx}(0,0) + (\sin\theta)^2 f_{yy}(0,0)]. \end{aligned}$$

Using (3.19), the normal curvature in direction $\mathbf{v} = [\cos\theta, \sin\theta, 0] \in T_O S$ is calculated as

$$k_n(O; \mathbf{v}) = k_n(O; \theta) = \frac{\mathbf{r}''(0) \cdot \boldsymbol{\nu}(O)}{|\mathbf{r}'(0)|^2} = (\cos\theta)^2 f_{xx}(0,0) + (\sin\theta)^2 f_{yy}(0,0). \tag{3.20}$$

3.2.1. *The principal curvatures and Euler's theorem.* We obtain a special version of what is called *Euler's theorem:* Remark that $k_1 = f_{xx}(0,0)$ is the normal curvature of S at O *in direction of the X-axis*, and $k_2 = f_{yy}(0,0)$ is the normal curvature of S at O *in direction of the Y-axis*.

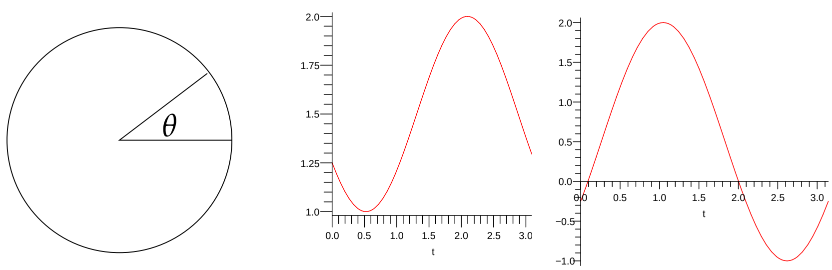
PROPOSITION 3.31. *Let S be the surface obtained as the* graph of a smooth function $f : D \to \mathbf{R}$, $D \subset \mathbf{R}^2$ *open with* $\mathbf{0} = [0,0] \in D$ *and*

$$f(0,0) = f_x(0,0) = f_y(0,0) = f_{xy}(0,0) = 0.$$

Then, $k_1 = f_{xx}(0,0)$ *and* $k_2 = f_{yy}(0,0)$ *are the* maximal, *resp.* minimal *normal curvatures of S at O, and the normal curvatures $k_n(\theta)$ in direction* $\mathbf{v} = [\cos\theta, \sin\theta, 0] \in T_O S$ *is given by* Euler's formula

$$k_n(O;\theta) = (\cos\theta)^2 k_1 + (\sin\theta)^2 k_2. \tag{3.21}$$

Euler's formula shows that the normal curvatures follow typical patterns as functions of the angle θ – as shown, at two different points, in Fig. 3.2.1 below:



PROOF:
Euler's formula (3.21) was already deduced as (3.20). It shows that $k_1 = k_n(O,0)$ and $k_2 = k_n(O,\frac{\pi}{2})$ are in fact normal curvatures. Why are they extremal? We may without restriction assume that $k_1 \geq k_2$. (If not, then exchange the role of the X- and the Y-axis). Then,

$$k_2 = ((\cos\theta)^2 + (\sin\theta)^2)k_2 \leq k_n(O;\theta) = (\cos\theta)^2 k_1 + (\sin\theta)^2 k_2 \leq$$
$$((\cos\theta)^2 + (\sin\theta)^2)k_1 = k_1.$$

□

DEFINITION 3.32. k_1 and k_2 are called the *principal curvatures* of S at O.

REMARK 3.33.
In particular, Euler's formula tells us how to calculate the normal curvatures in an *arbitrary* tangent direction if the principal curvatures are known. In order to find the normal curvature in a given tangent direction $\mathbf{v}$, it is enough to know the *two principal curvatures* and the *angle* between a *principal direction* – in our case the X-axis – and the direction given by $\mathbf{v}$. In fact, Euler's formula is true in the general case, too; this will be explained in Sec. 4.

EXAMPLE 3.34.
Let S be the surface obtained as the graph of the function $f : \mathbf{R}^2 \to \mathbf{R}$,
$f(x,y) = \cos y - \cos x$. Since $f_x(x,y) = \sin x$, $f_y(x,y) = -\sin y$ and $f_{xy}(x,y) = 0$, the surface S is of the type we have looked at so far. To calculate the principal curvatures, we need: $f_{xx}(x,y) = \cos x$ and $f_{yy}(x,y) = -\cos y$, and hence $k_1 = f_{xx}(0,0) = 1$,
$k_2 = f_{yy}(0,0) = -1$. By Euler's formula, the normal curvature at O in a (horizontal) direction with an angle θ with respect to the X-axis is $k_n(O;\theta) = (\cos\theta)^2 - (\sin\theta)^2$.

3.2.2. *The approximating paraboloid.* Formula (3.20) exhibits, that the normal curvature of S at O *only depends on the second derivatives* of f at O. Another way to phrase this is as follows:

DEFINITION 3.35. Let S be the graph of the function f above, let $k_1 = f_{xx}(0,0)$ and $k_2 = f_{yy}(0,0)$. Then, the second order Taylor approximation of f at $(0,0)$ takes the form

$$F(x,y) = k_1 x^2 + k_2 y^2.$$

The surface T given as the graph of the function F is called the *approximating paraboloid* of S at O.

EXAMPLE 3.36.
The approximating paraboloid to the graph surface S from Ex. 3.34 associated to
$f : \mathbf{R}^2 \to \mathbf{R},\ f(x,y) = \cos y - \cos x$ is the surface T associated to the function
$F : \mathbf{R}^2 \to \mathbf{R},\ F(x,y) = x^2 - y^2$, cf. Fig. 18.

In particular, the two surfaces S and T have the *same* principal curvatures and directions at O, and thus by Euler's formula (3.21), they have the same normal curvature in *every* tangent direction. And here is a parallel to the osculating circle along a curve - that contains all curvature information. Imagine an approximating paraboloid at *every* point of the surface. this so-called *osculating paraboloid* then contains the complete curvature information about the surface S, cf. Fig. 19.

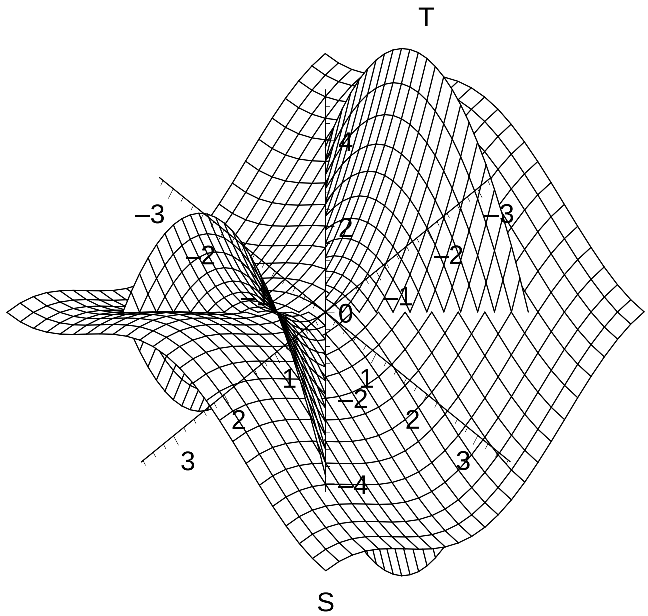


FIGURE 18. A surface with approximating paraboloid

4. Normal and geodesic curvature; the second fundamental form

In this chapter, we provide formulas for the normal curvatures at a given point of an *arbitrary* surface. The calculations are more involved than those in the preceeding special case. Along the way, we split the curvature of an *arbitrary* curve *on the surface S* into a *normal* and a *geodesic* component.

4.1. Normal and geodesic curvature. Let S denote a surface with a given parametrization

$$\mathbf{r} : \Omega \to \mathbf{R}^3,\ \Omega \subset \mathbf{R}^2.$$

Let C denote a curve *on* S with parametrization $\mathbf{x}(t) = \mathbf{r}(u(t), v(t))$ with $(u(t), v(t))$ a parametrization for the corresponding curve (cf. Def. 3.9) in Ω. Let P_t denote the point on the curve with $\overrightarrow{OP_t} = \mathbf{r}(u(t), v(t))$.

Along the curve, we have the following vector fields ("moving vectors"):

- the *tangent vector field* $\mathbf{t}(t)$ (or "moving tangent vector") attaching to each point on the curve the unit tangent vector.
- the *normal vectors to the surface* S – given by $\boldsymbol{\nu}(u,v) = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}(u,v)$ – assemble to a *normal vector field* ("moving normal vector") $\boldsymbol{\nu}(t) = \boldsymbol{\nu}(u(t), v(t))$ along the curve.
- The vector $\boldsymbol{\gamma}(t) = \boldsymbol{\nu}(t) \times \mathbf{t}(t)$ is contained in the tangent plane $T_{P_t}S$ at P_t and is perpendicular to $\mathbf{t}(t)$; it constitutes an "oriented normal vector" to $\mathbf{t}(t)$ with respect to $T_{P_t}S$".

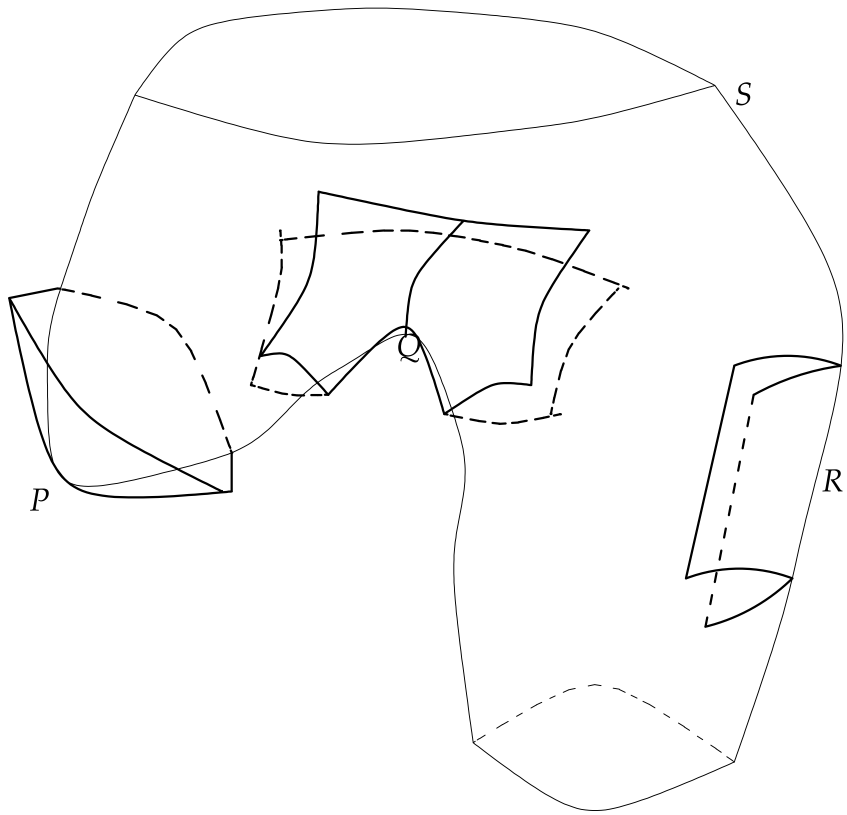


FIGURE 19. A surface with osculating paraboloid at several points

The three vectors $\mathbf{t}(t), \boldsymbol{\nu}(t), \boldsymbol{\gamma}(t)$ yield for us an alternative basis for $\mathbf{R}^3$ (a *moving frame*) along the curve.

The *principal normal vector* $\mathbf{n}(t)$ to the curve C at P_t is orthogonal to $\mathbf{t}(t)$ and hence a *unit* vector in the plane spanned by $\boldsymbol{\nu}(t)$ and $\boldsymbol{\gamma}(t)$. Thus, there is a variable angle $\theta(t)$ such that

$$\mathbf{n}(t) = \sin\theta(t)\boldsymbol{\gamma}(t) + \cos\theta(t)\boldsymbol{\nu}(t).$$

Hence, the *curvature vector* $\kappa(t)\mathbf{n}(t)$ splits into two components:

$$\kappa(t)\mathbf{n}(t) = \kappa(t)\sin\theta(t)\boldsymbol{\gamma}(t) + \kappa(t)\cos\theta(t)\boldsymbol{\nu}(t)$$

in a tangent and a normal direction.

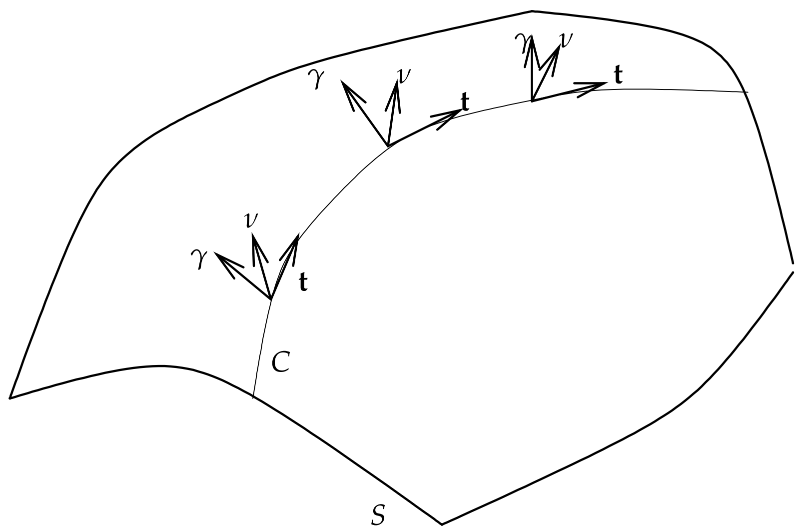


FIGURE 20. The moving frame along C

DEFINITION 3.37. The coefficients of these two components are called
- $k_g(t) = \kappa(t)\sin\theta(t)$ – the *geodesic curvature* at P_t;
- $k_n(t) = \kappa(t)\cos\theta(t) = \kappa(t)\mathbf{n}(t)\cdot\boldsymbol{\nu}(t)$ – the *normal curvature* at P_t.

REMARK 3.38.
In the case, that C is a *normal section* $C_{\mathbf{v}}(P)$ – which is contained in a normal *plane* $\pi_{\mathbf{v}}(P)$ – the principal normal vector $\mathbf{n}(t)$ to $C_{\mathbf{v}}(P)$ at P is a unit vector contained in that plane and perpendicular to $\mathbf{v}$. Hence, $\mathbf{n}(t)$ has to agree with $\boldsymbol{\nu}(t)$ or with $-\boldsymbol{\nu}(t)$, and therefore, the geodesic curvature of the normal section at P is $k_g(0) = 0$, whereas its normal curvature at P is $k_n(0) = \pm\kappa(0)$, i.e., the curvature of the normal section curve $C_{\mathbf{v}}(P)$. (The sign depends on the choice of the unit normal vector $\boldsymbol{\nu}(0)$.

The geodesic curvature of a curve will only be used later on. At this place, we want to calculate the normal curvature for a given curve C on the surface S using the parametrization $\mathbf{r}$ – and, in particular, the curvatures of normal sections: Let the curve be parametrized (in surface coordinates) as $\mathbf{x}(t) = \mathbf{r}(u(t), v(t))$. From Frenet's equations for a space curve (cf. Thm. 2.48), we know that

$$\mathbf{t}'(t) = s'(t)\kappa(t)\mathbf{n}(t) = s'(t)(k_g(t)\gamma(t) + k_n(t)\boldsymbol{\nu}(t)).$$

In particular,

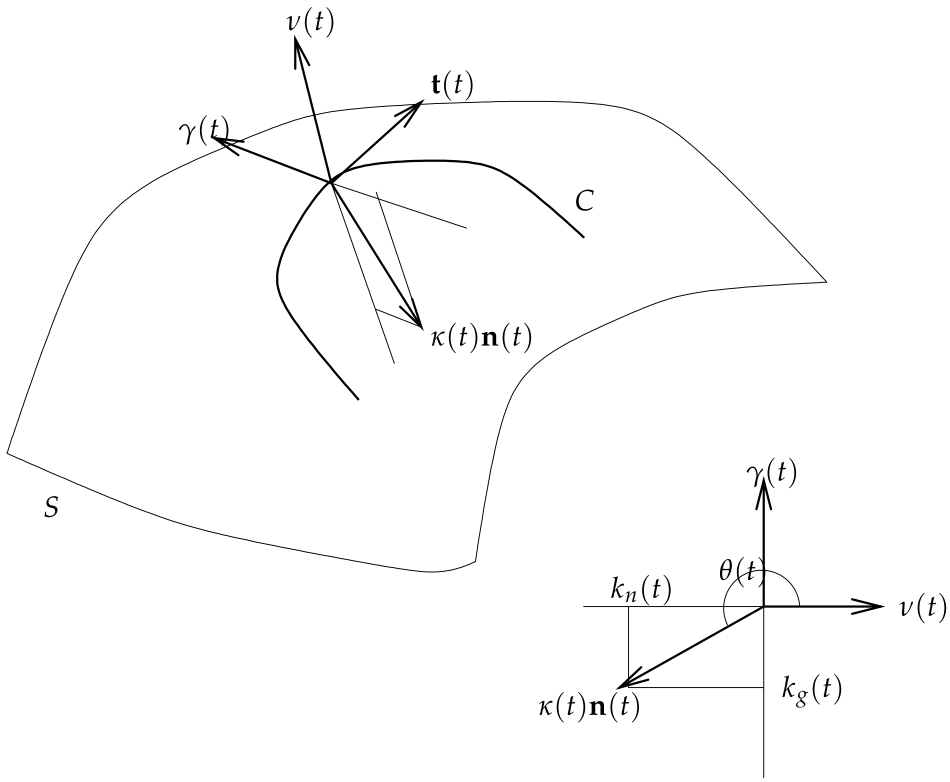


FIGURE 21. Decomposition of the curvature vector

PROPOSITION 3.39. *The normal curvature of k satisfies:*

$$k_n(t) = \frac{1}{s'(t)}(\mathbf{t}'(t) \cdot \boldsymbol{\nu}(t)) = -\frac{1}{s'(t)}(\mathbf{t}(t) \cdot \boldsymbol{\nu}'(t)). \tag{3.22}$$

PROOF:
The last equation follows by differentiating $\mathbf{t}(t) \cdot \nu(t) = 0$ (Remember the fundamental trick, Prop. 2.5.) □

To interpret the ingredients $\boldsymbol{\nu}' = (\boldsymbol{\nu} \circ \mathbf{x})'$ and $\mathbf{t}(t)$ in (3.22), we calculate them (along the curve C) using the chain rule (3.1), resp. the first of Frenet's equations (cf. Thm. 2.48), i.e.,

$$\begin{aligned} \boldsymbol{\nu}' &= \boldsymbol{\nu}_u u' + \boldsymbol{\nu}_v v' \\ \mathbf{t} &= \frac{1}{s'(t)}(\mathbf{r}_u u' + \mathbf{r}_v v'). \end{aligned}$$

We conclude from Prop. 3.39:

COROLLARY 3.40.

$$k_n(t) = -\frac{1}{s'(t)}\mathbf{t}(t)\cdot\boldsymbol{\nu}'(t)$$

$$= -\frac{1}{s'(t)^2}((\mathbf{r}_u\cdot\boldsymbol{\nu}_u)(u')^2 + (\mathbf{r}_u\cdot\boldsymbol{\nu}_v + \mathbf{r}_v\cdot\boldsymbol{\nu}_u)u'v' + (\mathbf{r}_v\cdot\boldsymbol{\nu}_v)(v')^2). \tag{3.23}$$

COROLLARY 3.41. Normal *curvature depends* only *on the tangent vector* **t** *of the curve at P, and* not *on the curve itself. For a* non-zero *tangent vector* $\mathbf{t} = a\mathbf{r}_u + b\mathbf{r}_v$*, the result of Cor.* 3.40 *can be used as the definition of the* normal curvature *of S* in direction **t***:*

$$k_n(\mathbf{t}) = -\frac{1}{|\mathbf{t}|^2}((\mathbf{r}_u\cdot\boldsymbol{\nu}_u)a^2 + (\mathbf{r}_u\cdot\boldsymbol{\nu}_v + \mathbf{r}_v\cdot\boldsymbol{\nu}_u)ab + (\mathbf{r}_v\cdot\boldsymbol{\nu}_v)b^2).$$

PROOF:
All you need in (3.23) resp. 3.41 is the normal vector $\boldsymbol{\nu}$ and its partial derivatives – depending on the surface and not on the curve – and the components $a = u'$ and $b = v'$ of the tangent vector **t**. □

4.2. The second fundamental form. Finally, we want to find expressions for the coefficients in the formula (3.23), that are easy to calculate. To this end, we use:

DEFINITION 3.42. Given a surface S with parametrization $r : D \to \mathbf{R}^3$ as above. We define three real-valued smooth functions $e, f, g : D \to \mathbf{R}$ by

$$\begin{aligned} e(u,v) &= (\mathbf{r}_{uu} \cdot \boldsymbol{\nu})(u,v) = \frac{\mathbf{r}_{uu} \cdot (\mathbf{r}_u \times \mathbf{r}_v)}{|\mathbf{r}_u \times \mathbf{r}_v|}(u,v); \\ f(u,v) &= (\mathbf{r}_{uv} \cdot \boldsymbol{\nu})(u,v) = \frac{\mathbf{r}_{uv} \cdot (\mathbf{r}_u \times \mathbf{r}_v)}{|\mathbf{r}_u \times \mathbf{r}_v|}(u,v); \\ g(u,v) &= (\mathbf{r}_{vv} \cdot \boldsymbol{\nu})(u,v) = \frac{\mathbf{r}_{vv} \cdot (\mathbf{r}_u \times \mathbf{r}_v)}{|\mathbf{r}_u \times \mathbf{r}_v|}(u,v), \end{aligned}$$

and the *second fundamental form* on a tangent vector $\mathbf{t} = a\mathbf{r}_u + b\mathbf{r}_v \in T_PS$ with $\overrightarrow{OP} = \mathbf{r}(u,v)$ as the quadratic polynomial in the two variables a and b

$$II(\mathbf{t})(u,v) = e(u,v)a^2 + 2f(u,v)ab + g(u,v)b^2.$$

The connection of these functions with the coefficients in (3.23) is explained in

LEMMA 3.43.

$$\begin{aligned} e(u,v) &= -(\mathbf{r}_u \cdot \boldsymbol{\nu}_{\boldsymbol{u}})(u,v); \\ f(u,v) &= -(\mathbf{r}_u \cdot \boldsymbol{\nu}_{\boldsymbol{v}})(u,v) = -(\mathbf{r}_v \cdot \boldsymbol{\nu}_{\boldsymbol{u}})(u,v); \\ g(u,v) &= -(\mathbf{r}_v \cdot \boldsymbol{\nu}_{\boldsymbol{v}})(u,v). \end{aligned}$$

PROOF:
Use partial differentiation of the equation

$$0 = \mathbf{r}_u \cdot \boldsymbol{\nu} = \mathbf{r}_v \cdot \boldsymbol{\nu}$$

with respect to both u and v. □

Since the length $|\mathbf{t}|$ of $\mathbf{t}$ is calculated by the first fundamental form (cf. Lemma 3.18), we get the following expression for the normal curvature in direction $\mathbf{t}$:

PROPOSITION 3.44. *The normal curvature k_n of S at P with $\overrightarrow{OP} = \mathbf{r}(u,v)$ in the tangent direction $\mathbf{t} = a\mathbf{r}_u + b\mathbf{r}_v \in T_PS$ is the quotient of the second and the first fundamental form (3.15) at $\mathbf{t}$, i.e.,*

$$k_n(\mathbf{t}) = \frac{II(\mathbf{t})}{I(\mathbf{t})} = \frac{e(u,v)a^2 + 2f(u,v)ab + g(u,v)b^2}{E(u,v)a^2 + 2F(u,v)ab + G(u,v)b^2}. \tag{3.24}$$

EXAMPLE 3.45.

(1) Let $\mathbf{r}(u,v) = [\rho\cos v, \rho\sin v, u]$ denote the parametrization for a cylinder from Ex. 3.6.1. We found earlier (cf. Ex. 3.12):

$$\mathbf{r}_u(u,v) = [0,0,1] \text{ and } \mathbf{r}_v(u,v) = [-\rho\sin v, \rho\cos v, 0],$$

and can thus calculate:

$$\mathbf{r}_{uu}(u,v) = [0,0,0],\ \mathbf{r}_{uv}(u,v) = [0,0,0] \text{ and } \mathbf{r}_{vv}(u,v) = [-\rho\cos v, -\rho\sin v, 0].$$

We calculated the normal vector in Ex. 3.15: $\boldsymbol{\nu}(u,v) = [-\cos v, -\sin v, 0]$. These ingredients allow us to calculate the coefficients of the second fundamental form from Def. 3.42:

$$\begin{aligned} e(u,v) &= (\mathbf{r}_{uu}\cdot\nu)(u,v) = \mathbf{0}\cdot[-\cos v, -\sin v, 0] = 0; \\ f(u,v) &= (\mathbf{r}_{uv}\cdot\nu)(u,v) = \mathbf{0}\cdot[-\cos v, -\sin v, 0] = 0; \\ g(u,v) &= (\mathbf{r}_{vv}\cdot\nu)(u,v) = [-\rho\cos v, -\rho\sin v, 0]\cdot[-\cos v, -\sin v, 0] \\ &= \rho((\cos v)^2 + (\sin v)^2) = \rho. \end{aligned}$$

The coefficients of the first fundamental form were already calculated in Ex. 3.22.1:

$$E(u,v) = 1, \quad F(u,v) = 0, \quad G(u,v) = \rho^2.$$

Hence, Prop. 3.44 allows us to calculate the normal curvature in direction $\mathbf{t} = a\mathbf{r}_u + b\mathbf{r}_v$ as

$$k_n(\mathbf{t}) = \frac{\rho b^2}{a^2 + \rho^2 b^2}.$$

In particular: $k_n(\mathbf{r}_u) = 0$, and $k_n(\mathbf{r}_v) = \frac{\rho}{\rho^2} = \frac{1}{\rho}$. This is as expected: These two numbers represent the normal curvature in direction of a straight line, resp. of a circle with radius ρ. All other normal curvatures are "sandwiched" between these two extrema (the principal curvatures):

$$0 \le k_n(\mathbf{t}) = \frac{\rho b^2}{a^2 + \rho^2 b^2} \le \frac{1}{\rho}.$$

(2) Similar calculations for a sphere $S(\rho)$ with radius ρ and parametrization from Ex. 3.6.2 yield:

$$e(u,v) = \rho(\cos v)^2 \quad f(u,v) = 0 \quad g(u,v) = \rho,$$
$$E(u,v) = \rho^2(\cos v)^2 \quad F(u,v) = 0 \quad G(u,v) = \rho^2.$$

We can calculate the normal curvature in direction $\mathbf{t} = a\mathbf{r}_u + b\mathbf{r}_v$ using Prop. 3.44:

$$k_n(\mathbf{t}) = \frac{\rho(\cos v)^2 a^2 + \rho b^2}{\rho^2(\cos v)^2 a^2 + \rho^2 b^2} = \frac{1}{\rho}\frac{\rho(\cos v)^2 a^2 + \rho b^2}{\rho(\cos v)^2 a^2 + \rho b^2} = \frac{1}{\rho},$$

i.e., the normal curvature takes the constant value $\frac{1}{\rho}$ in each tangent direction. Also this result is expected: Every normal section on a sphere is a (great) circle of radius ρ.

5. Principal curvatures, Gaussian curvature, and Mean curvature

5.1. Calculation of principal curvatures and principal directions. Our next aim is to generalise the result from Sect. 3.2.1 for the case of a graph surface to general surfaces. We hope to find at every point P on a surface S two *principal curvatures* such that all other normal curvatures $k_n(\mathbf{t})$ are sandwiched between those two. Moreover, it would be nice to have formulas calculating these entities. Our point of departure is Prop. 3.44 expressing normal curvatures as the quotient of the two fundamental forms on the tangent direction. What are the maximal, resp. minimal values for this expression (the normal curvature), and in which (tangent) directions do they occur?

Here is another way to phrase this question: Let $P \in S$ be such that $\overrightarrow{OP} = \mathbf{r}(u_0, v_0)$. We fix the values of the two fundamental forms *at that point*, i.e.,

$$E = E(u_0,v_0), \quad F = F(u_0,v_0), \quad G = G(u_0,v_0),$$
$$e = e(u_0,v_0), \quad f = f(u_0,v_0), \quad g = g(u_0,v_0).$$

Now we ask: For which real numbers k does the equation

$$k = \frac{ea^2 + 2fab + gb^2}{Ea^2 + 2Fab + Gb^2} \tag{3.25}$$

have a non-trivial solution $[a,b] \neq [0,0]$? (Remark: There exists always a solution for $k = \frac{e}{E}$, i.e., $[a,b] = [1,0]$.)

The following equations are equivalent to (3.25):

$$ea^2 + 2fab + gb^2 = kEa^2 + 2kFab + kGb^2,$$

$$(kE - e)a^2 + 2(kF - f)ab + (kG - g)b^2 = 0, \text{ and}$$

$$(kE - e)^2 a^2 + 2(kF - f)(kE - e)ab + (kG - g)(kE - e)b^2 = 0,$$

at least for $kE \neq e$. Using quadratic completion, this is equivalent to:

$$((kE - e)a + (kF - f)b)^2 + ((kG - g)(kE - e) - (kF - f)^2)b^2 = 0. \tag{3.26}$$

This equation can only have a solution if the second summand is less than or equal to zero, i.e., if

$$(kG-g)(kE-e)-(kF-f)^2 \leq 0, \text{ or} \tag{3.27}$$

$$(EG-F^2)k^2-(eG+gE-2fF)k+(eg-f^2) \leq 0, \text{ or}$$

$$k^2-\frac{eG+gE-2fF}{EG-F^2}k+\frac{eg-f^2}{EG-F^2} \leq 0. \tag{3.28}$$

(It is ok to divide by $EG-F^2$ since $EG-F^2>0$, cf. Rem. 3.27.)

DEFINITION 3.46. Let $P \in S$ be a point on a regular surface, and let E, F, G and e, f, g denote the coefficients of the 1. and second fundamental forms at P in a given parametrization. Then, we define the *Gaussian curvature*[a] $K(P)$ of S at P as the real number

$$K(p)=\frac{eg-f^2}{EG-F^2}, \tag{3.29}$$

and the *mean curvature* $H(P)$ as the real number

$$H(p)=\frac{eG+gE-2fF}{2(EG-F^2)}. \tag{3.30}$$

Remark, that K and H define smooth functions $K(u,v)$ and $H(u,v)$ on their domain.

[a]Here is a biography of the outstanding German mathematician and scientist Carl Friedrich Gauss

Using these abbreviations, the inequality (3.27) corresponds to

$$k^2-2Hk+K \leq 0.$$

For given real numbers K and H, the function $k \mapsto k^2-2Hk+K$ is quadratic; its graph is a parabola. Using (3.27), we observe, that the function has a *negative* value at $k=\frac{e}{E}$; at large positive or negative numbers k, it must have *positive* values. This is why the corresponding *equation*

$$k^2-2Hk+K=0$$

has two *real* solutions $k_1=H+\sqrt{H^2-K}$ and $k_2=H-\sqrt{H^2-K}$.

FIGURE 22. A stamp celebrating Carl Friedrich Gauss (1777 – 1855)

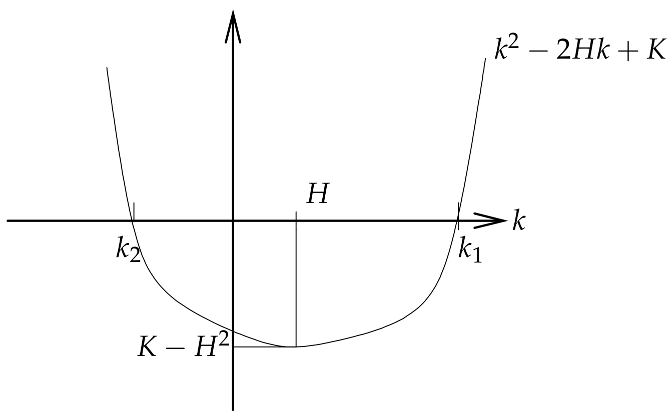


FIGURE 23. Graph of the function $k \mapsto k^2 - 2Hk + K$

DEFINITION 3.47. Let $P \in S$ be a point on a surface S, let $K(P)$ and $H(P)$ denote the Gaussian, resp. mean curvature of S at P. Then, the numbers

$$k_1(P) = H(P) + \sqrt{H(P)^2 - K(P)} \text{ and} \quad (3.31)$$

$$k_2(P) = H(P) - \sqrt{H(P)^2 - K(P)} \quad (3.32)$$

are called the *principal curvatures* for S at P.
The associated *principal directions* are the tangent directions $\mathbf{t}_1 = a_1\mathbf{r}_u + b_1\mathbf{r}_v$ and $\mathbf{t}_2 = a_2\mathbf{r}_u + b_2\mathbf{r}_v$ with $k_n(\mathbf{t}_i) = k_i(P)$, $i = 1, 2$. (These are only well-determined for $k_1 \neq k_2$!)

PROPOSITION 3.48. (1) *The principal curvatures, Gaussian curvature and mean curvature at a point $P \in S$ are connected by the following relations:*

$$K(P) = k_1(P)k_2(P)$$
$$H(P) = \frac{k_1(P)+k_2(P)}{2};$$

the last equation explains the name mean curvature*.*

(2) *The principal directions $\mathbf{t}_i = a_i\mathbf{r}_u + b_i\mathbf{r}_v$ can be determined as the solutions of the linear equations*

$$(k_iE-e)a_i + (k_iF-f)b_i = 0, \tag{3.33}$$

(*or as the solutions of the linear equations*

$$(k_iF-f)a_i + (k_iG-g)b_i = 0.)$$

PROOF:

(1) Calculate $k_1(P)k_2(P)$ and $k_1(P)+k_2(P)$ using the formulas from Def. 3.47.
(2) Have a look at equation (3.26):

$$((kE-e)a+(kF-f)b)^2 + ((kG-g)(kE-e)-(kF-f)^2)b^2 = 0.$$

The principal curvatures k_i satisfy the equation

$$(k_iG-g)(k_iE-e)-(k_iF-f)^2 = 0, \quad i = 1,2.$$

Hence, the first term

$$((k_iE-e)a+(k_iF-f)b)^2$$

has to vanish in the principal direction $\mathbf{t}_i$. This is equivalent to a solution $[a_i, b_i]$ of the linear equation

$$(k_iE-e)a_i+(k_iF-f)b_i = 0.$$

If you perform quadratic completion with respect to the b^2-term in (3.26), you arrive at the equation

$$((kE-e)(kG-g)-(kF-f)^2)a^2 + ((kF-f)a+(kG-g)b)^2 = 0.$$

The first term vanishes for $k = k_i$, so the second has to vanish for $[a,b] = [a_i,b_i]$, the coordinates of a principal direction.

□

REMARK 3.49.
Given a parametrization $\mathbf{r} : D \to \mathbf{R}^3$ of a surface S. Calculating principal curvatures and principal directions at a point $P \in S$ with $\overrightarrow{OP} = \mathbf{r}(u_0, v_0)$. you perform the following steps:

(1) Calculate the coefficients E, F, G and e, f, g of the first and of the second fundamental at (u_0, v_0) form using (3.10) and (3.24).
(2) Calculate the Gaussian curvature and the mean curvature at P using (3.29) and (3.30).
(3) Calculate $k_1(P)$ and $k_2(P)$ using (3.31).
(4) If $k_1(P) \neq k_2(P)$, calculate the principal directions $\mathbf{t}_1$ and $\mathbf{t}_2$ using (3.33).

EXAMPLE 3.50.

(1) Let $\mathbf{r}(u, v) = [\rho \cos v, \rho \sin v, u]$ denote the parametrization for a cylinder from Ex. 3.6.1. We found in Ex. 3.22 and Ex. 3.45:

$$\begin{array}{lll} E(u,v) = 1 & F(u,v) = 0 & G(u,v) = \rho^2; \\ e(u,v) = 0 & f(u,v) = 0 & g(u,v) = \rho. \end{array}$$

Hence, by (3.29) and (3.30) the Gauss and mean curvature take (constant) values

$$\begin{aligned} K &= \frac{eg - f^2}{EG - F^2} = \frac{0}{\rho^2} = 0; \\ H &= \frac{eG + gE - 2fF}{2(EG - F^2)} = \frac{\rho}{2\rho^2} = \frac{1}{2\rho}. \end{aligned}$$

Moreover, by (3.31), the principal curvatures at any point are:

$$k_1 = H + H = \frac{1}{\rho}, \qquad k_2 = H - H = 0.$$

Using (3.33), the principal directions are determined as solutions of the equations

$$\begin{array}{l} \frac{1}{\rho} a_1 = 0, \quad -\rho b_2 = 0, \text{ i.e.,} \\ \mathbf{t}_1 \in sp(\mathbf{r}_v), \quad \mathbf{t}_2 \in sp(\mathbf{r}_u). \end{array}$$

At P with $\overrightarrow{OP} = \mathbf{r}(u_0, v_0)$, we have thus:

The principal direction $\mathbf{t}_1 = [-\cos v, -\sin v, 0]$ is tangent to the horizontal circle on the cylinder through P, whereas the principal direction $\mathbf{t}_2 = [0, 0, 1]$ is tangent to the vertical line through P. Both have earlier been identified as normal sections. We have verified that they represent the normal sections with *maximal*, resp. *minimal* normal curvatures. Remark that the principal directions are perpendicular to each other.

(2) For the sphere $S(\rho)$ with radius ρ and parametrization from Ex. 3.6.2, we calculated in Ex. 3.45.2:

$$e(u,v) = \rho(\cos v)^2, \quad f(u,v) = 0, \quad g(u,v) = \rho,$$
$$E(u,v) = \rho^2(\cos v)^2, \quad F(u,v) = 0, \quad G(u,v) = \rho^2.$$

Hence, Gauss and mean curvature take the constant values

$$K = \frac{eg - f^2}{EG - F^2} = \frac{\rho^2(\cos v)^2}{\rho^4(\cos v)^2} = \frac{1}{\rho^2},$$
$$H = \frac{eG + gE - 2fF}{2(EG - F^2)} = \frac{2\rho^2(\cos v)^3}{2\rho^4(\cos v)^2} = \frac{1}{\rho},$$

and $k_1 = k_2 = \frac{1}{\rho}$. Since k_1 and k_2 are extremal and equal, *all* normal curvatures agree: $k_n(\mathbf{t}) = \frac{1}{\rho}$ for *every* tangent direction $\mathbf{t}$. For the same reason, every direction may be called a principal direction. This reflects rotational symmetry on a sphere.

Let us finally look at a special case in which calculations are particularly easy: Assume a surface has a parametrization such that at a given point $F = f = 0$. This situation arises, in particular at every point of a surface of revolution (cf. Sec. 6.2).

COROLLARY 3.51. *At a point P of a surface with $f = F = 0$, the princial curvatures and principal directions are given as*

$$k_1 = \frac{e}{E}; \quad \mathbf{t}_1 = \mathbf{r}_u; \quad k_2 = \frac{g}{G}; \quad \mathbf{t}_2 = \mathbf{r}_v.$$

PROOF:
At the point P, we have:

$$K(P) = \frac{eg}{EG} \text{ and } H(P) = \frac{eG + gE}{2EG}.$$

Using (3.31), we get the indicated values for the principal curvatures. The equations (3.33) determining principal directions read:

$$(\frac{gE}{G} - e)a_2 = 0; \quad (\frac{eG}{E} - g)b_1 = 0.$$

Hence, the principal directions are given by

$$\mathbf{t}_1 = \mathbf{r}_u(\text{for } b_1 = 0); \quad \mathbf{t}_2 = \mathbf{r}_v(\text{for } a_2 = 0).$$

□

5.2. The geometric significance of the Gaussian curvature. In Sect. 5.1, the Gaussian curvature appears just as a tool in the calculation of the principal curvatures. In fact, this invariant can tell us much more about the local and global properties of the surface S. First of all, one can see, that Gaussian curvature, mean curvature, and thus the principal curvatures are independent of the chosen parametrization – whereas the coefficients E, F, G and e, f, g clearly depend on parametrizations. The reason is, that the principal curvatures – as the extremal curvatures of the normal sections – are geometric entities that do not depend on parametrization. Gaussian curvature and mean curvature can be calculated from the principal curvatures, cf. Prop. 3.48.

5.2.1. *Classification of points on a surface.* Already the *sign* of the Gaussian curvature contains very useful information about the surface S in the neighbourhood of a given point $P \in S$:

DEFINITION 3.52. A point $P \in S$ is called
- *elliptic* if $K(P) > 0$,
- *hyperbolic* if $K(P) < 0$,
- *parabolic* if $K(P) = 0$ and $k_1(P) \neq 0$ or $k_2(P) \neq 0$,
- *planar* if $k_1(P) = k_2(P) = 0$.

The explanation for these names goes back to the relation

$$K(P) = k_1(P)k_2(P)$$

from Prop. 3.48. In particular, a point is
- *elliptic* if $k_1(P)$ and $k_2(P)$ have the *same* sign. Since $k_2(P) \leq k_n(\mathbf{t}) \leq k_1(P)$, all normal curvatures have then the same sign.
- *hyperbolic* if $k_1(P)$ and $k_2(P)$ have *different* signs, i.e., if $k_1(P) > 0$ and $k_2(P) < 0$; in particular, there is then a tangent direction $\mathbf{t}_0$ with $k_n(\mathbf{t}_0) = 0$. (This tangent direction $\mathbf{t}_0$ is called an *asymptotic* direction.)
- *parabolic* if one principal curvature is zero while the other is not.

Let us look at the different situations in some more detail. The (affine) tangent plane $\pi_p(S)$ (cf. Def. 3.10) divides Euclidean space into two half-spaces $\mathbf{E}^3_+$ and $\mathbf{E}^3_-$ containing the points P_+ with $\overrightarrow{PP_+} = \boldsymbol{\nu}$, resp. P_- with $\overrightarrow{PP_-} = -\boldsymbol{\nu}$.

elliptic point: Assume first $k_1(P) > k_2(P) > 0$. The curvature vector $\kappa\mathbf{n}$ of *every* normal section is then a positive multiple of the normal vector $\boldsymbol{\nu}(\boldsymbol{P})$ of the surface at P. Hence, *every* normal section at P bends into $\mathbf{E}^3_+$, i.e., contains a branch around P contained in $\mathbf{E}^3_+$. If $k_2(P) < k_1(P) < 0$, *every* normal section at P bends into $\mathbf{E}^3_-$. In conclusion, a neighbourhood of the surface close to P is contained in *one or the other* of the two half-spaces.

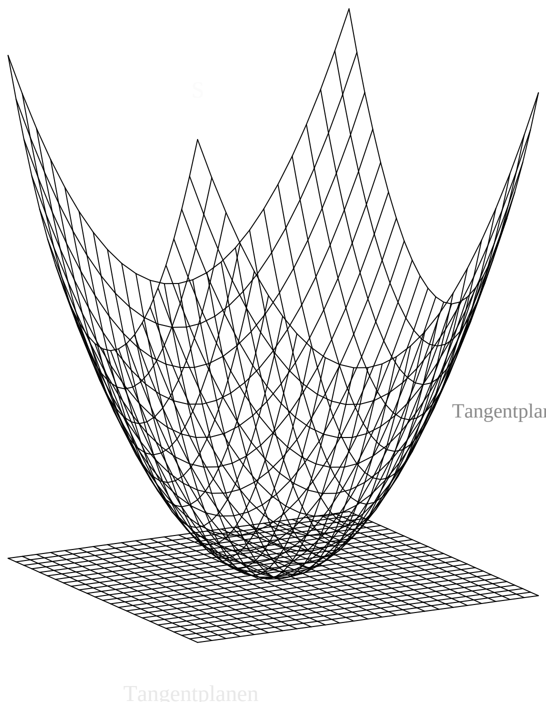


FIGURE 24. Elliptic point.

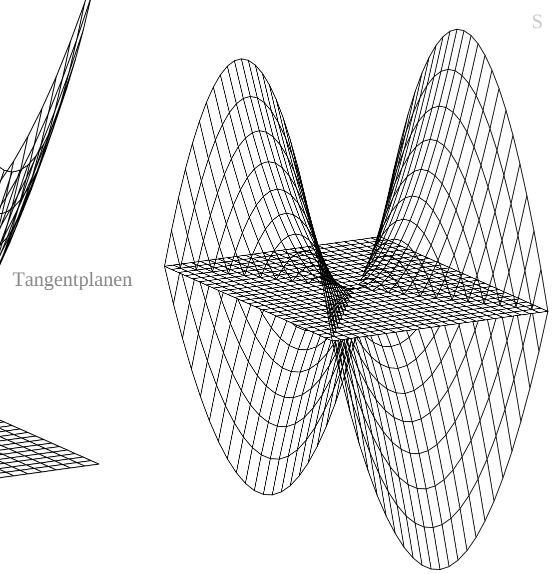


FIGURE 25. Hyperbolic Point.

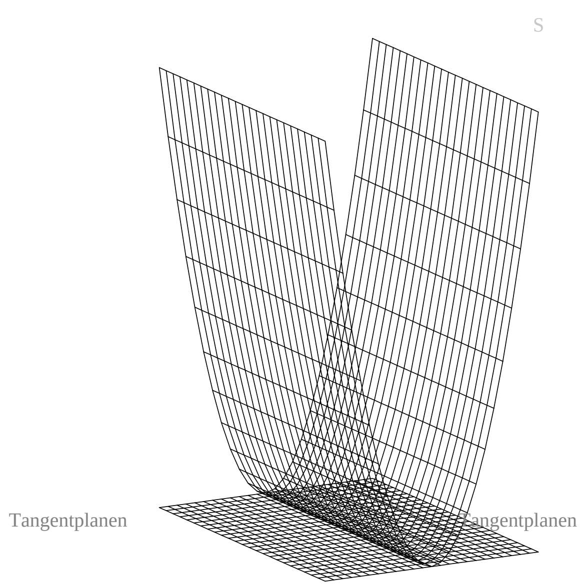


FIGURE 26. Parabolic point.

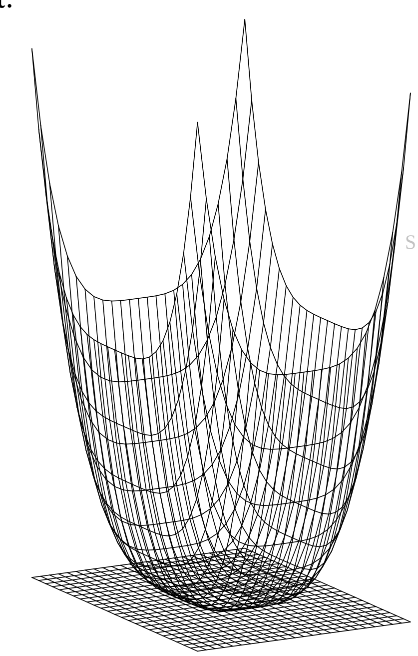


FIGURE 27. Planar Point.

hyperbolic point: The normal section in direction $\mathbf{t}_1$ bends into $\mathbf{E}^3_+$, while the normal section in direction $\mathbf{t}_2$ bends into $\mathbf{E}^3_-$. In conclusion, the affine tangent plane $\pi_p(S)$ cuts the surface into two pieces in a neighbourhood of the point P.

parabolic point: This is an intermediate situation. All normal curvatures are either greater than or equal to 0, resp. less than or equal to 0. That is, all normal sections bend into the same half-space with the possible exception of the principal direction with normal curvature 0.

planar point: *All* normal curvatures are zero. The affine tangent plane is the best 2nd order approximation of the surface near the point P, i.e., very close to S at a neighbourhood of P.

Another way to look at a surface S close to a point P is to describe the approximating paraboloid at P generalising Def. 3.35. To this end, we use a new coordinate system, cf. Sect. 4: Its origin is $O' = P$, its coordinate axes are given by $\mathbf{i}' = \mathbf{t}_1$, $\mathbf{j}' = \mathbf{t}_2$, and $\mathbf{k} = \mathbf{j} = \boldsymbol{\nu}$. (This is indeed an orthonormal coordinate system, since $\boldsymbol{\nu}$ is perpendicular to the tangent vectors $\mathbf{t}_1$ and $\mathbf{t}_2$, and, the principal directions $\mathbf{t}_1$ and $\mathbf{t}_2$ are orthogonal to each other, as we shall see in Sect. 5.3.) Using this coordinate system, we define

DEFINITION 3.53. The approximating paraboloid to the surface S at the point P is the graph of the function

$$z' = F(x', y') = k_1(x')^2 + k_2(y')^2$$

with k_1, k_2 the principal curvatures at P.

The approximating paraboloids are surfaces whose intersection with the planes $z' = z_0$ are conic sections, cf. Sect. 4.3, satisfying the equations $z_0 = k_1(x')^2 + k_2(y')^2$. These intersections are at a

- elliptic point — an ellipse or P or empty
- hyperbolic point — a hyperbola or two intersecting lines
- parabolic point — a parabola or P or empty
- planar point — the affine tangent plane or empty.

EXAMPLE 3.54.

Look at the surface of a torus (swimming belt), cf. Fig. 28. The points on the exterior are elliptic, those on the interior are hyperbolic, whereas the points on the top circle, and on the bottom circle are parabolic. A parametrization of the torus and the calculations showing the claims above will be given in Sect. 6.2.

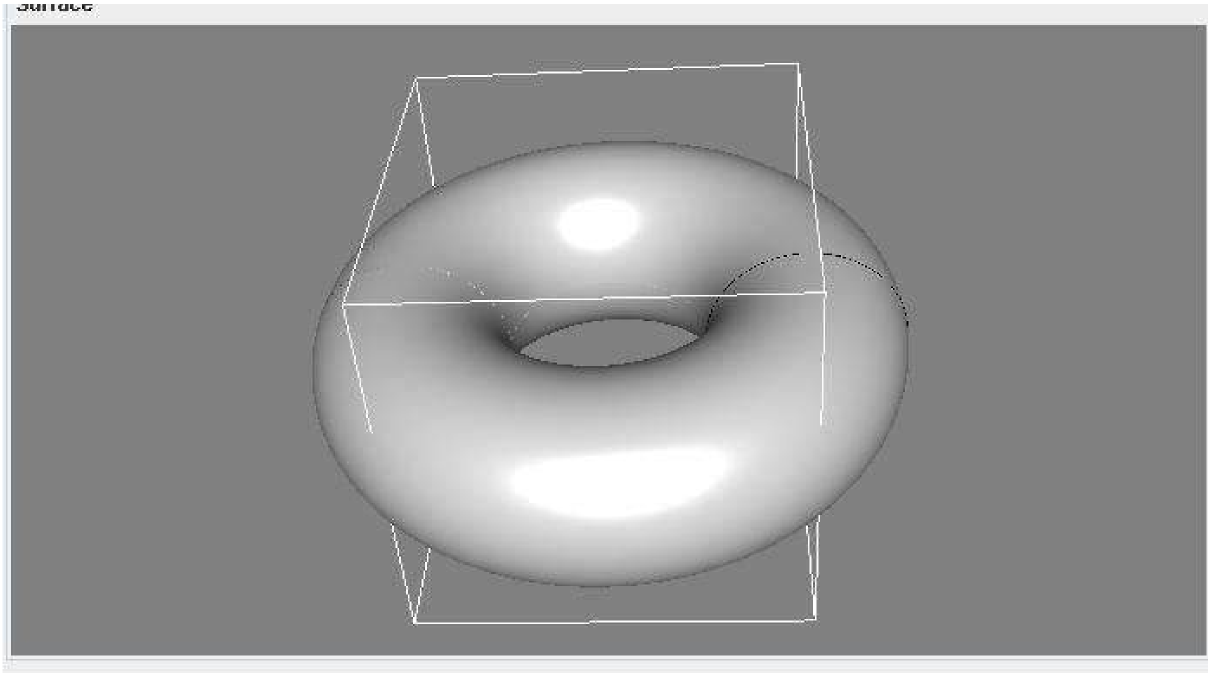


FIGURE 28. A torus (GEOLAB)

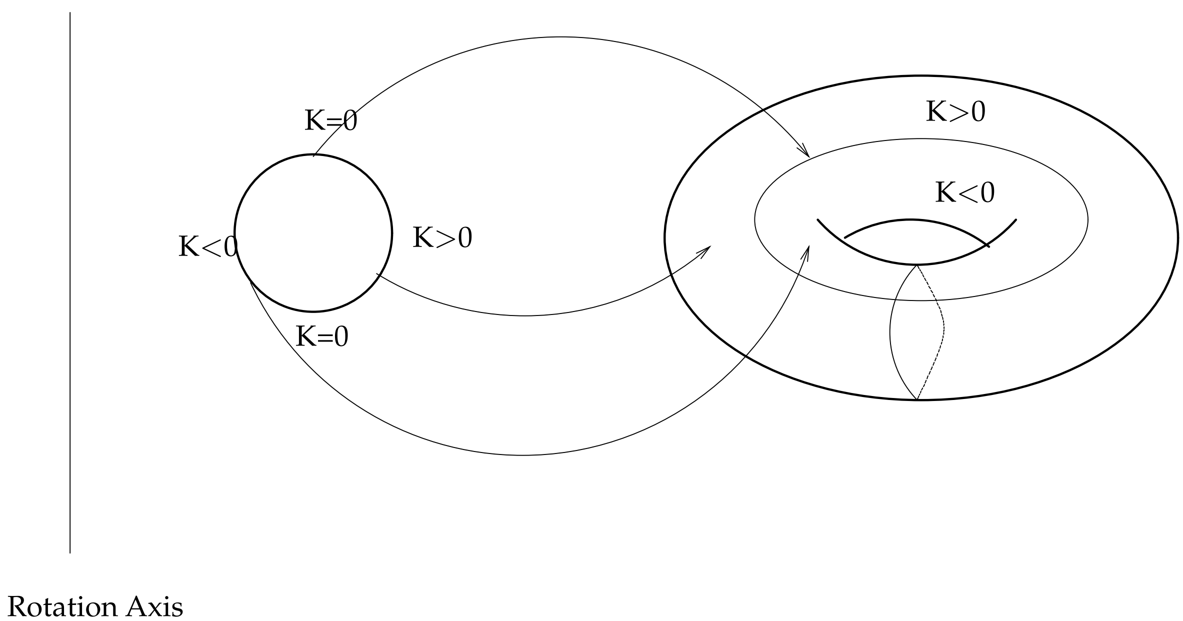


FIGURE 29. A torus contains elliptic, hyperbolic and parbolic points

5.2.2. *Gaussian curvature and isometries.* Many of the notions and discoveries in the differential geometry of surfaces go back to the great scientist C.F. Gauss. In his investigations of curvature properties (with a land surveying project as his point of departure), he was led to ask what happens to curvature when you *bend* a surface. For example, you may bend a sheet of paper (part of the plane) to get a cylinder. The plane has all normal curvatures zero. A cylinder has one zero principal curvature zero, while the other is non-zero, cf. Ex. 3.50. We conclude, that neither principal nor mean curvatures can be invariant under bendings, while the Gaussian curvature is invariant in this particular case. It would go beyond the framework of this course to explain the notions and to give a proof of the results of Gauss rigorously. But you should at least get a flavour of the ideas involved:

Let S_1 and S_2 denote two surfaces (e.g., a plane and a cylinder). Let $F : S_1 \to S_2$ denote a (differentiable) map from S_1 into S_2. To a curve C on S_1 (with parametrization $\mathbf{r} : I \to S_1$), we may associate the curve C_F on S_2 with parametrization $F \circ \mathbf{r} : I \to S_2$. An (invertible) map $F : S_1 \to S_2$ as above is called an *isometry* if and only if, for *every* curve C on S_1, the two curves C and C_F have the *same* length.

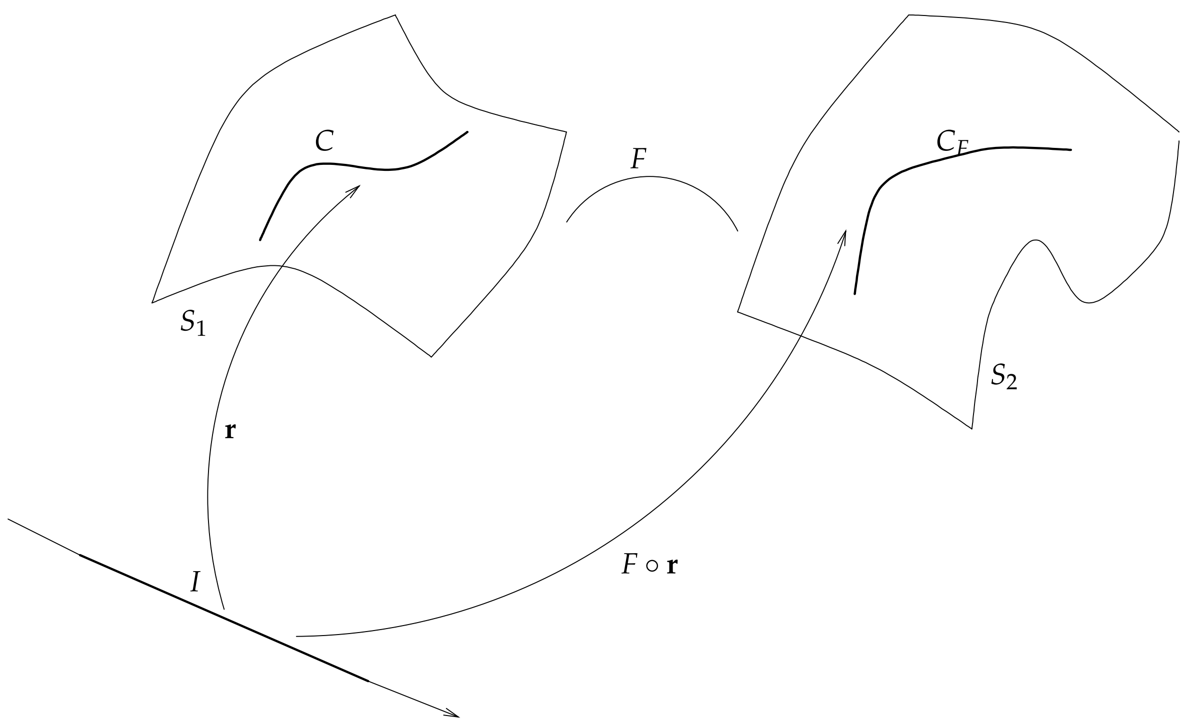


FIGURE 30. Curves C and C_F

Usually, one works with a more technical version of this definition: Curve length is calculated as the integral (over the interval I) of the velocity vectors corresponding to a parametrization. Thus one requires an isometry to preserve the length of tangent vectors; this is a condition that is easy to check.

EXAMPLE 3.55.

(1) Let $C_1 = \{[x,y,z] \in \mathbf{R}^3 | x^2+y^2=1\}$ denote the cylinder with radius 1 and the Z-axis as its axis. The map
$$F :]0, 2\pi[\times \mathbf{R} \to C_1, \quad F(u,v) = [\cos v, \sin v, u]$$
from a strip of size 2π in the plane to the cylinder is an isometry onto its image, i.e., it bends the strip around the cylinder. The length of any curve is preserved under the map F, cf. Fig. 31. (Technically speaking, F preserves the first fundamental form.)

(2) Let $S(1)$ denote a sphere of radius 1 centered at the origin; its north and south pole are denoted by N, resp. S. The map
$$G : S(1) \setminus \{N,S\} \to C_1, \quad [x,y,z] \to [\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2}, z]$$
projects the sphere from the axis containing N and S to the cylinder. (You may imagine a light bulb situated in the axis; obviously, there is no way to project the two poles). This map cannot be an isometry: A great circle (meridian) on the sphere connecting the two poles is projected to a straight line on the cylinder that is considerably shorter, cf.Fig. 31.

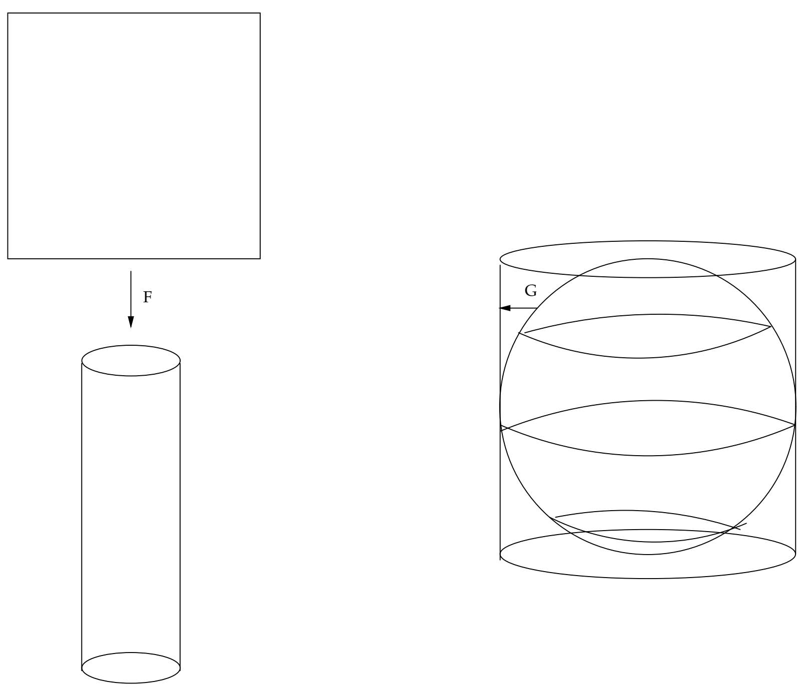


FIGURE 31. An isometry and a non-isometry

THEOREM 3.56. (***Theorema egregium, Gauss***)
An isometry preserves the Gaussian curvature, i.e.:
Let $F : S_1 \to S_2$ denote an isometry between two surfaces S_1 and S_2. For every $P \in S$, the Gaussian curvature at P and at $F(P) \in S_2$ agree:

$$K(P) = K(F(P)).$$

COROLLARY 3.57. *Let U denote any open set on a sphere $S(\rho)$ of radius ρ. There is no isometry between U and (any) open set in the plane.*

PROOF:
Assume there were such an isometry F. Then, for any $P \in U$,

$$\frac{1}{\rho^2} = K(P) = K(F(P)) = 0.$$

Contradiction! □

REMARK 3.58.
This argument shows that a map of a part of the earth with a fixed scale (1:1.000.000, say) does not exist! The maps you can buy indicating such a scale *must* have certain distortions. Those can be kept small at a minor scale (e.g., map of a city), but are very considerable for maps of the entire earth or of a continent.

5.3. Euler's Theorem. Let us finally show that the properties of principal curvatures, normal curvatures and principal directions that we considered for graphs in Sect. 3.2.1 are valid in general (cf. Fig. 3.2.1 for an illustration):

THEOREM 3.59. (***Euler's theorem)***[a] *Let S denote a surface, $P \in S$ a point with different principal curvatures $k_1 \neq k_2$. Let $\mathbf{t}_1, \mathbf{t}_2 \in T_P(S)$ denote unit vectors in the principal directions, and let $\mathbf{t} \in T_P(S)$ denote a unit tangent vector with an angle θ between $\mathbf{t}_1$ and $\mathbf{t}$. Then*

(1) *the principal directions are perpendicular to each other, i.e.,*

$$\mathbf{t}_1 \cdot \mathbf{t}_2 = 0;$$

(2) *The normal curvature $k_n(\mathbf{t})$ in direction $\mathbf{t}$ satisfies:*

$$k_n(\mathbf{t}) = (\cos\theta)^2 k_1 + (\sin\theta)^2 k_2$$

with k_i the principal curvature associated to $\mathbf{t}_i$. In particular,

$k_2 \leq k_n(\mathbf{t}) \leq k_1$ for every unit tangent vector $\mathbf{t} \in T_P(S)$.

[a]Here is a biography of the outstanding Swiss mathematician and scientist Leonhard Euler.

FIGURE 32. A stamp celebrating Leonhard Euler (1707 – 1783)

EXAMPLE 3.60.
Let $\mathbf{r}(u, v) = [\rho\cos v, \rho\sin v, u]$ denote the parametrization for a cylinder C_ρ from Ex. 3.6.1. We calculated the principal curvatures and principal directions associated to (any) point

P in Ex. 3.50: $k_1 = \frac{1}{\rho}$ in (horizontal) direction $\mathbf{t}_1 = \mathbf{r}_v$, and $k_2 = 0$ in (vertical) direction $\mathbf{t}_2 = \mathbf{r}_u$. Now let $\mathbf{t} = \cos\theta \mathbf{r}_v + \sin\theta \mathbf{r}_u$ denote an arbitrary tangent direction. The normal curvature in direction $\mathbf{t}$ is:

$$k_n(\mathbf{t}) = \frac{(\cos\theta)^2}{\rho^2}.$$

PROOF:
The proof below is quite computational and lengthy; a more conceptual proof is available using the theory of eigenvectors and eigenvalues.

(1) The equations (3.33) characterising the principal directions $\mathbf{t}_i = a_i\mathbf{r}_u + b_i\mathbf{r}_v$

$$\begin{aligned}(k_iE - e)a_i + (k_iF - f)b_i &= 0,\\ (k_iF - f)a_i + (k_iG - g)b_i &= 0.\end{aligned}$$

can be written down in matrix form as follows:

$$\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_i \\ b_i \end{bmatrix} = k_i\begin{bmatrix} E & F \\ F & G \end{bmatrix}\begin{bmatrix} a_i \\ b_i \end{bmatrix}.$$

We verify below $\mathbf{t}_1 \cdot \mathbf{t}_2 = 0$ by showing that $k_1(\mathbf{t}_1 \cdot \mathbf{t}_2) = k_2(\mathbf{t}_1 \cdot \mathbf{t}_2)$:

$k_1(\mathbf{t}_1 \cdot \mathbf{t}_2) = k_1(a_1\mathbf{r}_u + b_1\mathbf{r}_v) \cdot (a_2\mathbf{r}_u + b_2\mathbf{r}_v) =$

$= k_1[a_1, b_1]\begin{bmatrix} E & F \\ F & G \end{bmatrix}\begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = \left(k_1\begin{bmatrix} E & F \\ F & G \end{bmatrix}\begin{bmatrix} a_1 \\ b_1 \end{bmatrix}\right)^T\begin{bmatrix} a_2 \\ b_2 \end{bmatrix} =$

$= \left(\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_1 \\ b_1 \end{bmatrix}\right)^T\begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = [a_1, b_1]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_2 \\ b_2 \end{bmatrix} =$

$= [a_1, b_1]k_2\begin{bmatrix} E & F \\ F & G \end{bmatrix}\begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = k_2(a_1\mathbf{r}_u + b_1\mathbf{r}_v) \cdot (a_2\mathbf{r}_u + b_2\mathbf{r}_v) =$

$k_2(\mathbf{t}_1 \cdot \mathbf{t}_2)$.

(2) The vector $\mathbf{t} = \cos\theta\mathbf{t}_1 + \sin\theta\mathbf{t}_2$ is a unit vector, since
$|\mathbf{t}|^2 = \mathbf{t} \cdot \mathbf{t} = (\cos\theta\mathbf{t}_1 + \sin\theta\mathbf{t}_2) \cdot (\cos\theta\mathbf{t}_1 + \sin\theta\mathbf{t}_2) =$
$= (\cos\theta)^2(\mathbf{t}_1 \cdot \mathbf{t}_1) + (\sin\theta)^2(\mathbf{t}_2 \cdot \mathbf{t}_2) + 2\sin\theta\cos\theta(\mathbf{t}_1 \cdot \mathbf{t}_2) =$
$= (\cos\theta)^2 + (\sin\theta)^2 + 0 = 1.$

This is why $k_n(\mathbf{t}) = II(\mathbf{t}) =$

$= [a_1\cos\theta + a_2\sin\theta, b_1\cos\theta + b_2\sin\theta]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_1\cos\theta + a_2\sin\theta \\ b_1\cos\theta + b_2\sin\theta \end{bmatrix} =$

$= [a_1\cos\theta, b_1\cos\theta]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_1\cos\theta \\ b_1\cos\theta \end{bmatrix} + [a_1\cos\theta, b_1\cos\theta]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_2\sin\theta \\ b_2\sin\theta \end{bmatrix}$

$+ [a_2\sin\theta, b_2\sin\theta]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_1\cos\theta \\ b_1\cos\theta \end{bmatrix} + [a_2\sin\theta, b_2\sin\theta]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_2\sin\theta \\ b_2\sin\theta \end{bmatrix}$

$= (\cos\theta)^2[a_1, b_1]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} + \cos\theta\sin\theta[a_1, b_1]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$

$$+\sin\theta\cos\theta[a_2,b_2]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} + (\sin\theta)^2[a_2,b_2]\begin{bmatrix} e & f \\ f & g \end{bmatrix}\begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$$
$$= (\cos\theta)^2[a_1,b_1]k_1\begin{bmatrix} E & F \\ F & G \end{bmatrix}\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} + \cos\theta\sin\theta[a_1,b_1]k_2\begin{bmatrix} E & F \\ F & G \end{bmatrix}\begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$$
$$+\sin\theta\cos\theta[a_2,b_2]k_1\begin{bmatrix} E & F \\ F & G \end{bmatrix}\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} + (\sin\theta)^2[a_2,b_2]k_2\begin{bmatrix} E & F \\ F & G \end{bmatrix}\begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$$
$$= (\cos\theta)^2 k_1(\mathbf{t}_1\cdot\mathbf{t}_1) + \cos\theta\sin\theta k_2(\mathbf{t}_1\cdot\mathbf{t}_2) +$$
$$\sin\theta\cos\theta k_1(\mathbf{t}_2\cdot\mathbf{t}_1) + (\sin\theta)^2 k_2(\mathbf{t}_2\cdot\mathbf{t}_2)$$
$$= (\cos\theta)^2 k_1 + (\sin\theta)^2 k_2.$$

□

6. Special surfaces

Using the notions and formulas developped in the preceeding sections, we can now look at several classes of interesting surfaces, get new information about those, and shed new insight on the formalisms developped earlier.

6.1. Graph surfaces. Throughout this chapter, the graph S of a smooth function $f : D \to \mathbf{R},\ D \subset \mathbf{R}^2$ open, has served as one of our major examples. This is the place to collect many of the results previously obtained and to complete those with respect to curvature. As earlier, we shall use the parametrization

$$\mathbf{r} : D \to \mathbf{R}^3,\ \mathbf{r}(u,v) = [u,v,f(u,v)].$$

At the point $P \in S$ with $\overrightarrow{OP} = \mathbf{r}(u_0,v_0)$ and $(u_0,v_0) \in D$, we found in Ex. 3.22:

$$\mathbf{r}_u(u,v) = [1,0,f_u(u,v)] \text{ and } \mathbf{r}_v(u,v) = [0,1,f_v(u,v)],$$

and the coefficients of the first fundamental form (in terms of the prametrization $\mathbf{r}$):

$$E(u,v) = 1 + (f_u(u,v))^2;\ F(u,v) = f_u(u,v)f_v(u,v);\ G(u,v) = 1 + (f_v(u,v))^2.$$

To determine the coefficients of the second fundamental form, we need moreover:

$$\mathbf{r}_{uu}(u,v) = [0,0,f_{uu}(u,v)];\ \mathbf{r}_{uv}(u,v) = [0,0,f_{uv}(u,v)];\ \mathbf{r}_{vv}(u,v) = [0,0,f_{vv}(u,v)];$$

$$\boldsymbol{\nu}(u,v) = \frac{(\mathbf{r}_u \times \mathbf{r}_v)(u,v)}{|(\mathbf{r}_u \times \mathbf{r}_v)(u,v)|} = \frac{[-f_u(u,v), -f_v(u,v), 1]}{\sqrt{1 + (f_u(u,v))^2 + (f_v(u,v))^2}},$$

and obtain (Def. 3.42):

$$e(u,v) = \frac{f_{uu}(u,v)}{\sqrt{1 + (f_u(u,v))^2 + (f_v(u,v))^2}};\ f(u,v) = \frac{f_{uv}(u,v)}{\sqrt{1 + (f_u(u,v))^2 + (f_v(u,v))^2}};$$

$$g(u,v) = \frac{f_{vv}(u,v)}{\sqrt{1 + (f_u(u,v))^2 + (f_v(u,v))^2}}.$$

Furthermore, the Gauss and mean curvatures are:

$$K(u,v) = \frac{(f_{uu}f_{vv} - (f_{uv})^2)(u,v)}{(1 + f_u^2 + f_v^2)(u,v)};$$

$$H(u,v) = \frac{((1 + f_u^2)f_{vv} - 2f_u f_v f_{uv} + (1 + f_v^2)f_{vv})(u,v)}{2(1 + f_u^2 + f_v^2)^{\frac{3}{2}}(u,v)}.$$

Remark that the numerator above has the form of a *discriminant*. In fact, we can rediscover and rephrase the following result concerning the classification of critical points for a smooth function of two variables from Calculus:

PROPOSITION 3.61. *Let* $f : D \to \mathbf{R}$, $D \subset \mathbf{R}^2$ *open, denote a smooth function with critical point* $P_0 : (u_0, v_0) \in D$, *i.e.,* $f_u(u_0, v_0) = f_v(u_0, v_0) = 0$ – *both partial derivatives vanish. The function* f *has a local extremum* (*maximum or minimum*) *at* P_0 *if* $(f_{uu}f_{vv} - (f_{uv})^2)(u,v) > 0$, *whereas* P_0 *is a saddel point if* $(f_{uu}f_{vv} - (f_{uv})^2)(u,v) < 0$.

PROOF:
From Ex. 3.12.3, we see that the tangent plane to the graph of f at the critical point $[u_0, v_0, f(u_0, v_0)]$ is spanned by $[1,0,0]$ and $[0,1,0]$, hence horizontal and given by the equation $z = f(u_0, v_0)$.

The sign of the Gaussian curvature $K(u,v)$ and the sign of the discriminant $(f_{uu}f_{vv} - (f_{uv})^2)(u,v)$ agree. Therefore, we can use the results on the sign of the Gaussian curvature from Sect. 5.2: If $K(u,v) > 0$ – at an elliptic point – the surface "graph of f" is situated on *one* side of the horizontal tangent plane – at least locally – either over or under it; this is exactly the situation at a local extremum. If $K(u,v) < 0$ – at a hyperbolic point – there will be points on the graph on both sides of this horizontal plane close to the critical point; this is the situation at a saddel point. □

EXAMPLE 3.62.
Let $f : \mathbf{R}^2 \to \mathbf{R}$ denote the function $f(x,y) = \cos x + \cos y - 2$ with parametrization $\mathbf{r}(u,v) = [u, v, \cos u + \cos v - 2]$. Since $f_u(u,v) = -\sin u,\ f_v(u,v) = -\sin v,\ f_{uu}(u,v) = -\cos u,\ f_{uv}(u,v) = 0,$ $f_{vv}(u,v) = -\cos v$, we obtain:

$$E(u,v) = 1 + (\sin u)^2;\ F(u,v) = \sin u \sin v;\ G(u,v) = 1 + (\sin v)^2;$$

$$e(u,v) = \frac{-\cos u}{\sqrt{1 + (\sin u)^2 + (\sin v)^2}}\ f(u,v) = 0;\ g(u,v) = \frac{-\cos v}{\sqrt{1 + (\sin u)^2 + (\sin v)^2}};$$

$$K(u,v) = \frac{\cos u \cos v}{(1 + (\sin u)^2 + (\sin v)^2)^2};$$

$$H(u,v) = -\frac{\cos u(1 + (\sin u)^2) - \cos v(1 + (\sin v)^2)}{2(1 + (\sin u)^2 + (\sin v)^2)^{\frac{3}{2}}}.$$

The point $P_0 : (0,0)$ is *critical* with respect to the function f, i.e., both partial derivatives vanish at P_0. The other critical points have coordinates $(k\pi, l\pi)$ with integers k and l. At P_0, the graph of f has the XY-plane $z = 0$ as its tangent plane; at the other critical points the tangent plane is one of the horizontal planes $z = 0$ or $z = -2$ or $z = -4$. The Gaussian curvature at P_0 is $K(0,0) = 1$, the mean curvature is $H(0,0) = -1$. The principal curvatures at this point are $k_1(P_0) = k_2(P_0) = -1$. This means, that all directions are principal directions. The approximating paraboloid at P_0 is a sphere of radius 1. Since the Gaussian curvature is positive, the function has a local extremum (in fact, a local maximum) at P_0.

At $(k\pi, l\pi)$, the Gaussian curvature is positive, if k and l are either both even or both odd. The function f has a local maxiumum at such a point ($f(k\pi, l\pi) = 0$) if k and l are both even and a local minimum ($f(k\pi, l\pi) = -4$) if k and l are both odd. If k is even and l is odd – or vice versa – , then $K(k\pi, l\pi) = -1$, and f has a saddel point at $(k\pi, l\pi)$ with horizontal tangent plane $z = -2$.

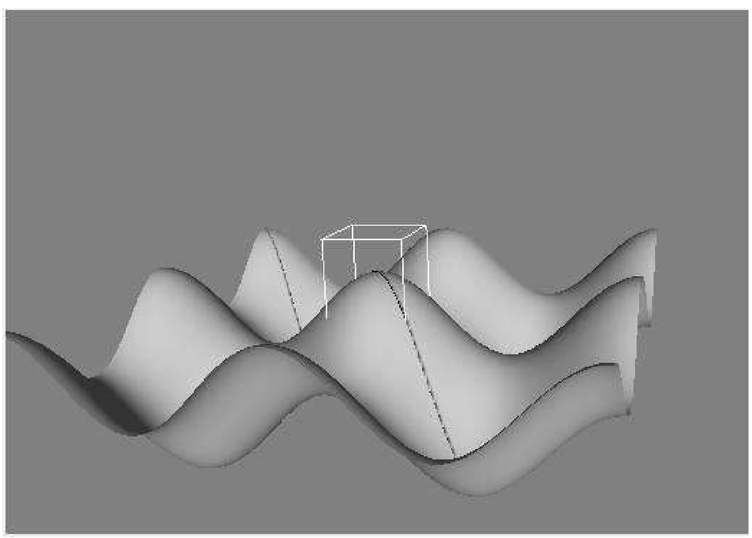


FIGURE 33. Graph of the function $z = \cos x + \cos y - 2$

6.2. Surfaces of revolution. Let $P : [x, 0, z] \in \mathbf{E}^3$ denote a point in the XZ-plane in Euclidean space. Rotating it around the Z-axis one obtains a circle with radius u

in the plane $Z = z$. This circle can be parameterised in cylindrical coordinates (polar coordinates instead of XY-variables, Z-coordinate unchanged) by the vector function

$$\mathbf{r}(u) = [x \cos u, x \sin u, z],\ u \in [0, 2\pi].$$

6.2.1. *Generating a surface of revolution.* Instead of a single point, one may just as well rotate *all* the points on a regular simple curve (i.e., without self-intersections) in the XZ-plane around the Z-axis. Such a (*generating*) curve with a smooth parametrisation $\mathbf{x}(v) = [x(v), 0, z(v)]$ gives then rise to the parametrisation

(3.34) $$\mathbf{r}(u, v) = [x(v) \cos u, x(v) \sin u, z(v)],\ a < v < b,\ 0 < u < 2\pi.$$

The effect is, that the whole curve is rotated around the Z-axis and generates a *surface of revolution*. We have to be a bit careful for several reasons:

(1) If the function $x(v)$ takes both negative and positive values, it has to take a value $x(v_0) = 0$ somewhere. This corresponds to a point $[0, 0, z(v_0)]$ on the Z-axis, which stays *fixed* under the rotation. The surface of revolution thus obtained would *not* be regular in this point. This is why we require

(3.35) $$x(v) > 0 \text{ for all } v > 0.$$

(2) In order to get an *open* domain for the parametrization, we have insisted on the constraint $0 < u < 2\pi$. Hence, the parametrization does *not* cover the entire surface we have in mind (one point is lacking for every circle it consists of). To cover the whole surface, one can use the same formula for the parametrization, but within the interval $u \in]-\pi, \pi[$ in the domain.

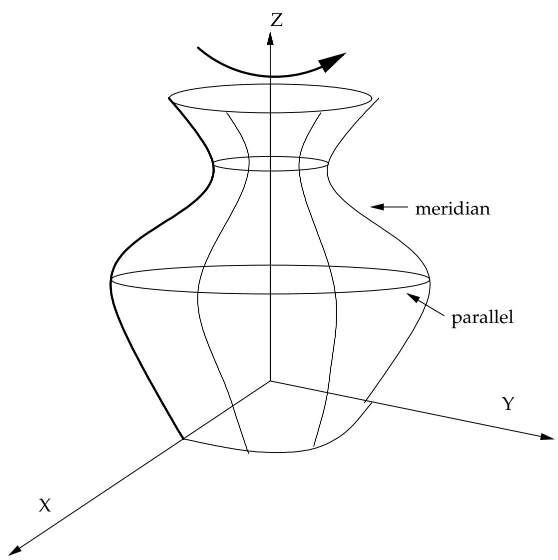


FIGURE 34. Generating curve, surface of revolution, parallels, and meridians

EXAMPLE 3.63.

(1) Rotating a vertical line segment in the XZ-plane around the Z-axis generates a cylinder; rotating a skew line segment yields part of a cone.
(2) Rotating a half circle with center on the Z-axis (but without the points on the Z-axis gives rise to a sphere (without the poles on the Z-axis); rotating a circle around the Z-axis contained in the XZ-plane yields a *torus* (swimming belt) cf. Ex. 3.54 and Sect. 6.2.3.

6.2.2. *Properties of a surface of revolution.* Let us now analyse the properties of a surface of revolution using the parametrization (3.34). The parameter curves of the surface are

- the *parallels* with parametrization $[x(v_0)\cos u, x(v_0)\sin u, z(v_0)]$; those are circles for every fixed parameter v_0; and
- the *meridians* (rotated copies of the generating curve) with parametrization $[x(v)\cos u_0, x(v)\sin u_0, z(v)]$ for a fixed parameter u_0.

Let us now check that the vector function (3.34) satisfies the requirements to a parametrization from Def. 3.4. Since the generating curve is simple, and since the domain is chosen with an open interval $u \in]0, 2\pi[$, it is easy to check, that $\mathbf{r}$ is one-to-one. The partial derivatives of $\mathbf{r}$ are calculated as

$$\mathbf{r}_u(u,v) = x(v)[-\sin u, \cos u, 0]; \quad \mathbf{r}_v(u,v) = [x'(v)\cos u, x'(v)\sin u, z'(v)].$$

To see that these two vectors are linearly independent for every parameter pair (u,v), we calculate their cross product

$$(\mathbf{r}_u \times \mathbf{r}_v)(u,v) = x(v)[z'(v)\cos u, z'(v)\sin u, -x'(v)],$$

and show that it cannot be $\mathbf{0}$. By (3.35), $x(v) > 0$. Moreover, the vector $[z'(v)\cos u, z'(v)\sin u, -x'(v)]$ has length $\sqrt{z'(v)^2 + x'(v)^2}$. This is exactly the speed along the generating (regular) curve, and thus nonzero everywhere.

The normal vector at $P_{uv} \in S$ with $\overrightarrow{OP_{uv}} = \mathbf{r}(u,v)$ is

$$\boldsymbol{\nu}(u,v) = \frac{1}{\sqrt{z'(v)^2 + x'(v)^2}}[z'(v)\cos u, z'(v)\sin u, -x'(v)].$$

Check also the calculation of the second partial derivatives of $\mathbf{r}$:

$$\mathbf{r}_{uu}(u,v) = -x(v)[\cos u, \sin u, 0]; \quad \mathbf{r}_{uv}(u,v) = x'(v)[-\sin u, \cos u, 0];$$
$$\mathbf{r}_{vv}(u,v) = [x''(v)\cos u, x''(v)\sin u, z''(v)].$$

Let us now calculate the coefficients of the two fundamental forms (cf. Def. 3.16 and Def. 3.42):

$$E(u,v) = x^2(v); \quad F(u,v) = 0; \quad G(u,v) = x'(v)^2 + z'(v)^2;$$
$$e(u,v) = -\frac{x(v)z'(v)}{\sqrt{x'(v)^2 + z'(v)^2}}; \quad f(u,v) = 0; \quad g(u,v) = \frac{x''(v)z'(v) - x'(v)z''(v)}{\sqrt{x'(v)^2 + z'(v)^2}}.$$

Since $F = 0$, we conclude that *the parallels and the meridians are orthogonal to each other* at every point of the surface. Moreover, since $F = f = 0$, we can use Cor. 3.51 to determine principal curvatures and directions: The tangents $\mathbf{r}_u$ to the parallel circles, resp. $\mathbf{r}_v$ to the meridians are the *principal directions*. The principal curvatures are:

$$k_1 = \frac{e}{E} = \frac{-z'(v)}{x(v)\sqrt{x'(v)^2 + z'(v)^2}}; \quad k_2 = \frac{g}{G} = \frac{x''(v)z'(v) - x'(v)z''(v)}{(x'(v)^2 + z'(v)^2)^{\frac{3}{2}}}.$$

To get closer to the geometric meaning and also in order to simplify, let us assume that the generating curve is parametrised by arc length. In this case, the speed $\sqrt{x'(v)^2 + z'(v)^2}$ corresponding to the parametrization takes the constant value 1 at every parameter v. In particular, the principal curvatures simplify to:

$$k_1 = \frac{-z'(v)}{x(v)}; \quad k_2 = x''(v)z'(v) - x'(v)z''(v).$$

COROLLARY 3.64. *Let C denote a generating curve* parameterized by arc length. *Let S denote the associated surface of revolution, and $P_{uv} \in S$ the point with $\overrightarrow{OP_{uv}} = \mathbf{r}(u, v)$. The Gauss curvature at P_{uv} is given as*

$$K(P_{uv}) = -\frac{x''(v)}{x(v)}. \tag{3.36}$$

In particular, P_{uv} is elliptic if $x''(v) < 0$, hyperbolic, if $x''(v) > 0$, and parabolic or plane if $x''(v) = 0$.

PROOF:
We differentiate the relation: "1=(speed)$^2 = x'(v)^2 + z'(v)^2$" with respect to v and obtain: $0 = x'(v)x''(v) + z'(v)z''(v)$. Hence,

$$K(P_{uv}) = k_1(u, v)k_2(u, v) = -\frac{z'(v)(x''(v)z'(v) - x'(v)z''(v))}{x(v)} =$$

$$= -\frac{z'(v)^2 x''(v) + x'(v)^2 x''(v)}{x(v)} = -\frac{x''(v)}{x(v)}.$$

□

REMARK 3.65.
Formula (3.36) can be applied to obtain parametrizations of surfaces of revolution with *constant* Gaussian curvature K. All you have to do is to solve the system of differential equations:

$$\begin{aligned} x''(v) - Kx(v) &= 0; \\ x'(v)^2 + y'(v)^2 &= 1. \end{aligned}$$

The solutions include the spheres with radius $K^{-\frac{1}{2}}$ for $K > 0$, the plane for $K = 0$, but also a variety of different (non-bounded) surfaces, in fact even for $K < 0$. The latter serve as model spaces for *hyperbolic* geometry, which probably describes the large scale properties of our universe much more accurate than Euclidean geometry.

6.2.3. *A torus.* Finally, let us generate a surface of revolution by rotating a circle with radius r and center on the X-axis at $[R, 0, 0]$ with $R > r$, cf. Figures 28, 29 and 35. The parametrization by arc length of this circle is given by

$$\mathbf{x}(v) = [R + r\cos\frac{v}{r}, 0, r\sin\frac{v}{r}], \quad 0 \le v \le 2\pi r,$$

cf. Ex. 2.24. In this case, both the parallels and the meridians of the resulting *torus* are circles; their tangents indicate the principal directions. Since $x(v) = R + r\cos\frac{v}{r}$, we obtain from (3.36) the following formula for every point P_{uv} on the parallel circle corresponding to the parameter value v:

$$K(P_{uv}) = -\frac{x''(v)}{x(v)} = \frac{\cos\frac{v}{r}}{r(R + r\cos\frac{v}{r})}.$$

By assumption, the denominator is always positive. Indeed, we have verified the results of Ex. 3.54 by computation: The point P_{uv} is classified by the signature of the expression $\cos\frac{v}{r}$ in the numerator. If it is positive (for $v < \frac{\pi}{2}r$ or $v > \frac{3\pi}{2}r$), the point P_{uv} is situated on the exterior and elliptic; if it is negative (for $\frac{\pi}{2}r < v < \frac{3\pi}{2}r$), the point P_{uv} is situated on the interior and hyperbolic. It is zero for $v = \frac{\pi}{2}r$ and for $v = \frac{3\pi}{2}r$ corresponding to the parabolic points on the top, resp. bottom circles.

6.3. Ruled surfaces. At a first thought it might seem strange that a curved surface can be made up of straight line segments. But we have already met such a strange subject: The helicoid from Ex. 3.8.2 is generated by horizontal lines connecting a helix and an axis. The helicoid is thus an example of a *ruled* surface, generated by moving a straight line (segment) in Euclidean space. Other easy examples are cylinders and cones. These are moreover examples of *developable* surfaces, which are locally isometric to the plane. In particular, they can be constructed by *bending* or *rolling* a (plane) sheet of metal.

In general, a ruled surface has a parametrization of the form

$$\mathbf{r}(u, v) = \mathbf{p}(u) + v\mathbf{q}(u), \quad u \in]a, b[, \quad v \in]c, d[. \tag{3.37}$$

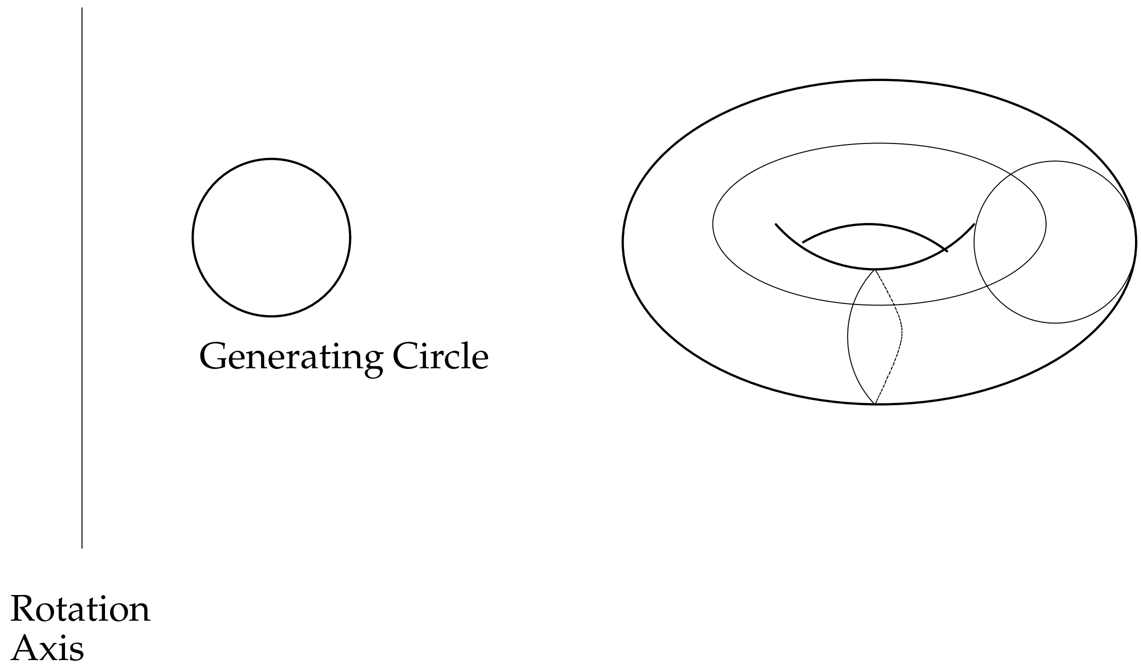


FIGURE 35. A torus as a surface of revolution

For $v = 0$, we obtain the curve given by $\mathbf{p}(u),\ u \in]a,b[$, as one of the 1. parameter curves ; it is called the *directrix* of the surface; the 2. parameter curves (fixing a value $u = u_0$) are *line segments*; they are called the *rulings* of the surface.

EXAMPLE 3.66.

(1) Let $\mathbf{p}(u) = \mathbf{p}_0 \in \mathbf{R}^3$ be *constant*. The ruled surface given by (3.37) (for $c > 0$) is then a (generalised) *cone*.
(2) If $\mathbf{p}(u)$ is a parametrization for a *plane* curve, and $\mathbf{q}(u) = \mathbf{q}_0 \in \mathbf{R}^3$ is *constant*, the ruled surface given by (3.37) is a (generalised) *cylinder*.
(3) If $\mathbf{q}(u) = \mathbf{p}'(u),\ u \in]a,b[$, the ruled surface given by (3.37) is a *tangent surface*. It consists of all tangent lines to the directrix.
(4) For a constant $c > 0$, the ruled surface with $\mathbf{p}(u) = [u,0,0]$ and $\mathbf{q}(u) = [0,1,cu]$ has parametrisation $\mathbf{r}(u,v) = [u,v,cuv]$. It consists of the points $[x,y,z] \in \mathbf{E}^3$ satisfying the equation $z = cxy$. It represents a *hyperbolic paraboloid* consisting of layers of horizontal hyperbola (for $z = z_0$).

See Fig. 36 for illustrations.

What do we know about the Gaussian curvature of a ruled surface right away? Well, there is a straight line (with curvature 0) through each point, and its (tangent) direction must correspond to *normal curvature* 0. By Euler's Theorem (Thm. 3.59), this value 0 is

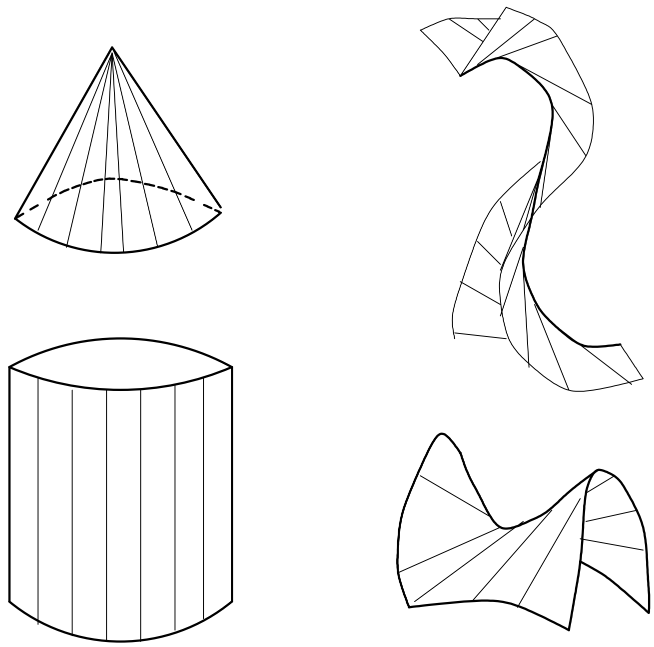


FIGURE 36. Cone, cylinder, tangent surface, and hyperbolic paraboloid

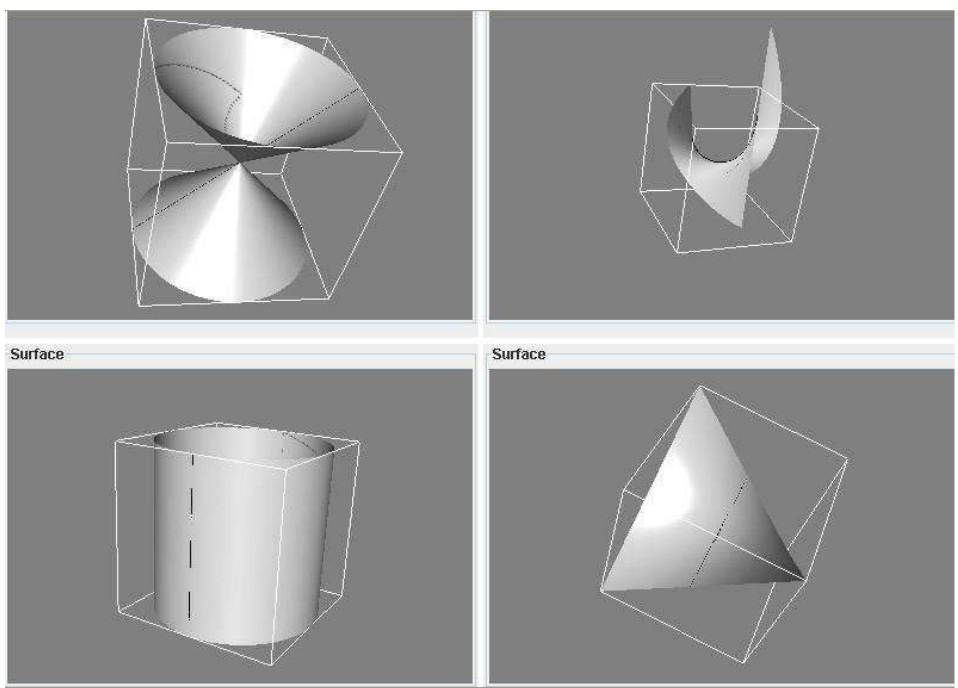


FIGURE 37. The same ruled surfaces as above

sandwiched between the two principal curvatures: $k_2 \leq 0 \leq k_1$. Hence, the princpal curvatures cannot have the same sign, and *the Gaussian curvature K has to be less than or equal to zero everywhere*.

In fact, the two possibilities $K = 0$, resp. $K < 0$, correspond to two classes of ruled surfaces: To see this, we calculate some of the important geometric entities using the parametrization (3.37):

$$\mathbf{r}_u(u,v) = \mathbf{p}'(u) + v\mathbf{q}'(u); \quad \mathbf{r}_v(u,v) = \mathbf{q}(u);$$
$$(\mathbf{r}_u \times \mathbf{r}_v)(u,v) = \mathbf{p}'(u) \times \mathbf{q}(u) + v\mathbf{q}'(u) \times \mathbf{q}(u);$$
$$\mathbf{r}_{uv}(u,v) = \mathbf{q}'(u); \quad \mathbf{r}_{vv}(u,v) = 0.$$

$$f(u,v) = \frac{\mathbf{q}'(u) \cdot (\mathbf{p}'(u) \times \mathbf{q}(u))}{|(\mathbf{r}_u \times \mathbf{r}_v)(u,v)|}; g(u,v) = 0. \tag{3.38}$$

In particular, the Gaussian curvature $K = \frac{-f^2}{EG-F^2} \leq 0$, and

$$K(u,v) = 0 \Leftrightarrow f(u,v) = 0 \Leftrightarrow D(u) = [\mathbf{p}'(u), \mathbf{q}(u), \mathbf{q}'(u)] = \mathbf{q}'(u) \cdot (\mathbf{p}'(u) \times \mathbf{q}(u)) = 0.$$

Remark, that a ruled surface S with parametrization (3.37) in general is not regular at every point $P \in S$. In the example of a tangent surface, the points on the curve generating it are *singular*, as you might guess from Fig. 36. One can show for general *non-cylindrical* ruled surfaces (i.e., $\mathbf{q}'(u) \neq \mathbf{0}$ for all $u \in]a,b[$), that the singular points are always situated on a particular curve, the *line of striction*.

PROPOSITION 3.67. *Let S denote a ruled surface with parametrization*

$$\mathbf{r}(u,v) = \mathbf{p}(u) + v\mathbf{q}(u), \quad u \in]a,b[, \quad v \in]c,d[.$$

The tangent planes $T_{P_{u_0,v}}S$ along a particular ruling corresponding to $u = u_0$ are

- *constant* $\Leftrightarrow K(u_0,0) = 0$;
- *run through* all *linear planes in $\mathbf{R}^3$ containing $sp(\mathbf{q}(u_0))$ except $sp(\mathbf{q}(u_0), \mathbf{q}'(u_0))$* $\Leftrightarrow K(u_0,0) \neq 0$.

PROOF:
The tangent plane $T_{P_{u_0,v}}S$ is the plane spanned by the vectors $\mathbf{r}_u(u_0,v)$ and $\mathbf{r}_v(u_0,v)$, i.e., $T_{P_{u_0,v}}S = sp(\mathbf{p}'(u_0) + v\mathbf{q}'(u_0), \mathbf{q}(u_0))$. It is always contained in the subspace spanned by the three vectors $\mathbf{q}(u_0), \mathbf{q}'(u_0), \mathbf{p}'(u_0) \in \mathbf{R}^3$. By (3.38) and the following remark, $K(u_0,0) = 0 \Leftrightarrow$ the space product (cf. Sect. 1.2.4)
$D(u_0) = [\mathbf{q}(u_0), \mathbf{q}'(u_0), \mathbf{p}'(u_0)] = \mathbf{q}'(u_0) \cdot (\mathbf{p}'(u_0) \times \mathbf{q}(u_0)) = 0$. In particular,
$K(u_0,0) = 0$ if and only if the three vectors $\mathbf{q}(u_0), \mathbf{q}'(u_0), \mathbf{p}'(u_0)$ are *linearly dependent*.

In this case, they only span a linear *plane*, and hence $T_{P_{u_0,v}}S = sp(\mathbf{p}'(u_0) + v\mathbf{q}'(u_0)) = sp(\mathbf{q}(u_0), \mathbf{q}'(u_0), \mathbf{p}'(u_0))$ has to be equal to this plane independent of the value of v. For $K(u_0, 0) \neq 0$, the three vectors span the entire vector space $\mathbf{R}^3$. For $v \in \mathbf{R}$, the vectors $\mathbf{p}'(u_0) + v\mathbf{q}'(u_0)$ hit every linear line in $sp(\mathbf{p}'(u_0), \mathbf{q}'(u_0))$ except the one through $\mathbf{q}'(u_0)$. □

DEFINITION 3.68. A ruled surface with parametrization

$$\mathbf{r}(u,v) = \mathbf{p}(u) + v\mathbf{q}(u), \quad u \in]a,b[, \quad v \in]c,d[$$

and $|\mathbf{q}(u)| = 1$ for all u is called *developable* if the space product (cf. Sect. 1.2.4)

$$D(u) = [\mathbf{q}(u), \mathbf{q}'(u), \mathbf{p}'(u)] = 0 \text{ for every } u \in]a,b[.$$

COROLLARY 3.69. *A developable surface S has Gaussian curvature 0 at all regular points $P \in S$. By a theorem of Minding, S is locally isometric* (cf. Sect. 5.2.2) *to the plane.*

EXAMPLE 3.70.
Let S denote the hyperbolic paraboloid from Ex. 3.66.4. with $\mathbf{p}(u) = [u, 0, 0]$ and $\mathbf{q}(u) = [0, 1, cu]$, and hence $\mathbf{p}'(u) = [1, 0, 0]$ and $\mathbf{q}'(u) = [0, 0, c]$. In this case, the space product

$$D(u) = [\mathbf{q}(u), \mathbf{q}'(u), \mathbf{p}'(u)] = \begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ cu & c & 0 \end{vmatrix} = c \neq 0.$$

In particular, the tangent planes are *not* constant along any ruling, and the paraboloid is *not* developable.

7. The geometric labotatory for surfaces

A prototype of a geometric laboratory for *surfaces* is under development. It aims to provide applets giving the user the possibility to plot surfaces using one or several parametrizations, to show tangent planes, normal vectors etc. in a dynamical framework. The first applets have been developed by Martin Quist, Aalborg University and can be reached at www.math.aau.dk/~raussen/VIDIGEO/GEOLAB/SURFACE. You can specify a parametrization by a formula $X(u,v) =, Y(u,v) =, Z(u,v) =$, you can specify the range of (u,v) (at least as a rectangle), you can visualize the image of a curve $(u(t), v(t))$ etc. Rotation of the surface is obtained by clicking on the *left* mouse button; you zoom in and out using the middle mouse button.

The applet in Fig. 38 illustrates a parametrization. It allows you to follow the image of a point and/or of a curve in the parameter plane on the surface itself.

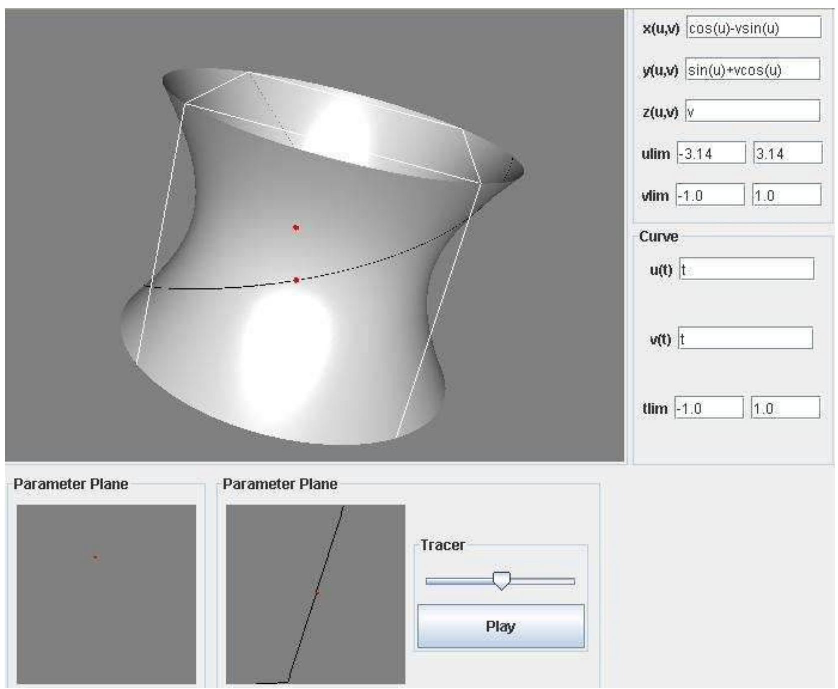


FIGURE 38. Illustration of a parametrization

Index