Unit 1. Basic Structures on $\mathbb{R}^n$, Length of Curves.

$\mathbb{R}^n$, addition of vectors and multiplication by scalars, vector spaces over $\mathbb{R}$, linear combinations, linear independence, basis, dimension, linear and affine linear subspaces, tangent space at a point, tangent bundle; dot product, length of vectors, the standard metric on $\mathbb{R}^n$; balls, open subsets, the standard topology on $\mathbb{R}^n$, continuous maps and homeomorphisms; simple arcs and parameterized continuous curves, reparameterization, length of curves, integral formula for differentiable curves, parameterization by arc length.

THE LINEAR STRUCTURE OF $\mathbb{R}^n$

Recall that $\mathbb{R}^n$ is the set of n-tuples of real numbers

$$\mathbb{R}^n = \{ (x_1,\dots,x_n) : x_i \in \mathbb{R} \}.$$

We shall denote the elements of $\mathbb{R}^n$ by underlined letters.

We know from 3-dimensional geometry that the introduction of a Cartesian coordinate system in space yields an identification between points, vectors and 3-tuples of real numbers (each point can be identified by its position vector and its coordinates). By analogy, we can think of the elements of $\mathbb{R}^n$ (which are just n-tuples of reals) as points of an n-dimensional space or vectors.

If $\mathbf{x} = (x_1,\dots,x_n)$ and $\mathbf{y} = (y_1,\dots,y_n)$ are two elements of $\mathbb{R}^n$ and $\lambda \in \mathbb{R}$ is a real number then we define the sum and difference of $\mathbf{x}$ and $\mathbf{y}$ and the scalar multiple of $\mathbf{x}$ by

$$\mathbf{x} \pm \mathbf{y} = (x_1 \pm y_1,\dots,x_n \pm y_n)$$

$$\lambda\,\mathbf{x} = (\lambda\, x_1,\dots,\lambda\, x_n).$$

$\mathbb{R}^n$ is a vector space over the field of real numbers with respect to these operations.

(Let us recall the definition of an abstract vector space. A set X is a vector space over the field of real numbers if X is equipped with a binary operation +, and for each $\lambda \in \mathbb{R}$ the multiplication of elements of X with λ is defined in such a way that the following identities are satisfied

$((\mathbf{x} + \mathbf{y}) + \mathbf{z}) = (\mathbf{x} + (\mathbf{y} + \mathbf{z}))$ (associativity);

$\exists\, \mathbf{0} \in X$ such that $\mathbf{x} + \mathbf{0} = \mathbf{x}$ for all $\mathbf{x} \in X$;

$\forall\, \mathbf{x} \in X\ \exists\, -\mathbf{x} \in X$ such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$;

$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ (commutativity);

$\lambda(\mathbf{x} + \mathbf{y}) = \lambda\,\mathbf{x} + \lambda\,\mathbf{y}$;

$(\lambda + \mu)\,\mathbf{x} = \lambda\,\mathbf{x} + \mu\,\mathbf{x}$;

$(\lambda\,\mu)\,\mathbf{x} = \lambda\,(\,\mu\,\mathbf{x}\,)$;

$1\,\mathbf{x} = \mathbf{x}$.

Exercise: Check that $\mathbb{R}^n$ is indeed a vector space.

A linear combination of some vectors $\mathbf{x}_1,\ldots,\mathbf{x}_k \in X$ is a vector of the form

$$\lambda_1\mathbf{x}_1 + \ldots + \lambda_k\mathbf{x}_k,$$

where the λ_i's are real numbers. The vectors $\mathbf{x}_1,\ldots,\mathbf{x}_k \in X$ are linearly independent if their linear combination can be $\mathbf{0}$ only if all coefficients λ_i are 0. A basis of X is a maximal set of linearly independent vectors. The dimension of the vector space X is the cardinality of a basis.)

Exercise: Show that $\mathbb{R}^n$ is n-dimensional.

A non-empty subset V of $\mathbb{R}^n$ is a linear subspace if for all $\mathbf{x}$, $\mathbf{y} \in V$ and $\lambda \in \mathbb{R}$ we have $\mathbf{x} + \mathbf{y} \in V$ and $\lambda\,\mathbf{x} \in V$. Linear subspaces of $\mathbb{R}^n$ are vector spaces themselves, so their dimension is defined. A 0-dimensional subspace consists of the single point $\mathbf{0}$. 1-dimensional linear subspaces are straight lines passing through the origin, 2-dimensional linear subspaces are ordinary planes that go through the origin etc. Linear subspaces always go through the origin.

A subset of $\mathbb{R}^n$ is an affine linear subspace if it is a translate of a linear subspace. The dimension of an affine linear subspace is the dimension of the linear subspace from which it is obtained by a translation. k-dimensional affine linear subspaces of $\mathbb{R}^n$ will be called shortly k-planes. 0-planes are points, 1-planes are straight lines, 2-planes are the ordinary 2-dimensional planes of $\mathbb{R}^n$. (n-1)-planes of $\mathbb{R}^n$ are also called hyperplanes.

In 3-dimensional geometry, a vector is an equivalence class of directed segments, where two directed segments represent the same vector if they have the same direction and length. Given a point $\mathbf{p}$ and a vector $\mathbf{x}$, there is a unique directed segment that starts from $\mathbf{p}$ and represents $\mathbf{x}$. Returning to $\mathbb{R}^n$, we shall call a pair $(\mathbf{p},\mathbf{x}) \in \mathbb{R}^n \times \mathbb{R}^n$ a tangent vector of $\mathbb{R}^n$ at $\mathbf{p}$ or a vector based at $\mathbf{p}$ (and we may think of this pair as a directed segment that initiates from $\mathbf{p}$ and represents the vector $\mathbf{x}$). The set of all tangent vectors of $\mathbb{R}^n$ at $\mathbf{p}$ form the tangent space of $\mathbb{R}^n$ at $\mathbf{p}$ and is denoted by $T_{\mathbf{p}}\mathbb{R}^n$. There is a one to one correspondence between $\mathbb{R}^n$ and $T_{\mathbf{p}}\mathbb{R}^n$ established by the "forgetting" mapping

$$T_{\mathbf{p}}\mathbb{R}^n \longrightarrow \mathbb{R}^n \qquad (\mathbf{p},\mathbf{x}) \longmapsto \mathbf{x},$$

with the help of which we can furnish $T_{\mathbf{p}}\mathbb{R}^n$ with a vector space structure in a natural way.

The tangent bundle $T_*\mathbb{R}^n$ of $\mathbb{R}^n$ is the disjoint union of all tangent spaces of $\mathbb{R}^n$: $T_*\mathbb{R}^n = \bigcup_{\mathbf{p}\in\mathbb{R}^n} T_{\mathbf{p}}\mathbb{R}^n$. The mapping π: $T_*\mathbb{R}^n \longrightarrow \mathbb{R}^n$, $(\mathbf{p},\mathbf{x}) \longmapsto \mathbf{p}$ that assigns to each tangent vector the point at which it is tangent is called the canonical projection of the tangent bundle.

THE METRIC OF $\mathbb{R}^n$

The linear structure of $\mathbb{R}^n$ is not enough to measure distance, area, etc. To measure e.g. the length of curves we need a metric. The standard metric on $\mathbb{R}^n$ is connected with a bilinear function called the dot product. The dot product of the vectors $\mathbf{x} = (x_1,\dots,x_n)$ and $\mathbf{y} = (y_1,\dots,y_n)$ is defined by the formula

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i .$$

With the help of the dot product one can define the norm or length of a vector and the distance of two points as follows

$$\| \mathbf{x} \| = \sqrt{\langle \mathbf{x} , \mathbf{x} \rangle} , \quad d(\mathbf{x},\mathbf{y}) = \| \mathbf{x} - \mathbf{y} \| .$$

Proposition. We have

i) $d(\mathbf{x},\mathbf{y}) \geqslant 0$, and equality holds if and only if $\mathbf{x} = \mathbf{y}$;

ii) $d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$;

iii) $d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z}) \geqslant d(\mathbf{x},\mathbf{z})$ (triangle inequality)

for any $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in \mathbb{R}^n$.

Proof. It is enough to prove iii), the rest is obvious. Introducing the vectors $\mathbf{a} = \mathbf{y} - \mathbf{x}$ and $\mathbf{b} = \mathbf{z} - \mathbf{y}$, iii) is equivalent to the inequalities

$$\| \mathbf{a} \| + \| \mathbf{b} \| \geqslant \| \mathbf{a} + \mathbf{b} \| ,$$

$$\| \mathbf{a} \|^2 + \| \mathbf{b} \|^2 + 2\| \mathbf{a} \| \| \mathbf{b} \| \geqslant \| \mathbf{a} \|^2 + \| \mathbf{b} \|^2 + 2 \langle \mathbf{a},\mathbf{b} \rangle ,$$

$$\| \mathbf{a} \| \| \mathbf{b} \| \geqslant \langle \mathbf{a},\mathbf{b} \rangle .$$

The last inequality is known as the Cauchy-Schwartz-Bunyakovsky inequality and can be proved as follows. Consider the second degree polynomial of the variable λ defined by the equality

$$p(\lambda) = (\mathbf{a} + \lambda\mathbf{b})^2 = \lambda^2\| \mathbf{b} \|^2 + \lambda\, 2\langle \mathbf{a},\mathbf{b} \rangle + \| \mathbf{a} \|^2 .$$

Since we have $p(\lambda) \geqslant 0$ for all λ, the discriminant of p must be non-positive, i.e.

$$0 \geqslant 4\langle \mathbf{a},\mathbf{b}\rangle^2 - 4\,\|\mathbf{a}\|^2\,\|\mathbf{b}\|^2,$$

and this is just what we wanted to prove. ▮

Remark. If X is a set and d is a function defined on the direct product X x X, such that conditions i), ii) and iii) are satisfied, then d is called a metric on X and the pair (X , d) is called a metric space.

THE TOPOLOGY OF $\mathbb{R}^n$

Definition. The open ball in $\mathbb{R}^n$ with center $\mathbf{x} \in \mathbb{R}^n$ and radius $\varepsilon > 0$ (or an ε-ball centered at $\mathbf{x}$) is the set $B_\varepsilon(\mathbf{x}) = \{ \mathbf{y} \in \mathbb{R}^n : d(\mathbf{x},\mathbf{y}) < \varepsilon \}$. Similarly, the closed ball in $\mathbb{R}^n$ with center $\mathbf{x} \in \mathbb{R}^n$ and radius $\varepsilon > 0$ (or a closed ε-ball centered at $\mathbf{x}$) is defined as the set $\overline{B}_\varepsilon(\mathbf{x})=\{\mathbf{y} \in \mathbb{R}^n : d(\mathbf{x},\mathbf{y}) \leqslant \varepsilon\}$.

Definition. A subset U of $\mathbb{R}^n$ is called open if for each $\mathbf{x} \in U$ there is a positive ε such that the ball $B_\varepsilon(\mathbf{x})$ is contained in U.

Definition A pair (X,τ) is said to be a topological space if X is a set, and τ is a family of subsets of X, called the open subsets of X, such that

i) the empty set $\emptyset$ and X are open;

ii) the intersection of two open subsets is also open;

iii) the union of an arbitrary family of open subsets is open.

The family τ of open subsets is called the topology on X.

Examples.

a) Let X be an arbitrary set. The discrete topology on X is the "maximal topology" on X, in which every subset is open.

b) The antidiscrete topology on X is the "minimal topology" on X, in which only the empty set and X are open.

c) The standard topology of $\mathbb{R}^n$.

Proposition. The family of open subsets defines a topology on $\mathbb{R}^n$. This topology is referred to as the standard topology on $\mathbb{R}^n$.

Proof. Obviously, the empty set and $\mathbb{R}^n$ are open. If U and V are open subsets, and $\mathbf{x}$ is a common point of theirs, then there exist positive numbers ε_1, ε_2 such that $B_{\varepsilon_1}(\mathbf{x}) \subset U$ and $B_{\varepsilon_2}(\mathbf{x}) \subset V$. Let ε be the minimum of ε_1 and ε_2. Then $B_\varepsilon(\mathbf{x}) \subset U \cap V$ showing that the intersection $U \cap V$ is open. Finally, let $\{U_i : i \in I\}$ be an arbitrary family of open sets, and $\mathbf{x}$ be an element of their union. Then we can find an index $j \in I$ and a positive ε such that $B_\varepsilon(\mathbf{x}) \subset U_j \subset \bigcup_{i\in I} U_i$, thus the union $\bigcup_{i\in I} U_i$ is open. ▮

d) The topology of a topological space defines a topology on every of its subsets by the following construction.

Proposition. Let Y be a subset of a topological space (X, τ). Then the family $\tau|_Y = \{ U \cap Y : U \in \tau \}$ defines a topology on Y.

The proof of the proposition is straightforward from the following identities.

i) $\emptyset \cap Y = \emptyset, \qquad X \cap Y = Y$;

ii) $(U \cap Y) \cap (V \cap Y) = (U \cap V) \cap Y$;

iii) $\bigcup_{i \in I} (U_i \cap Y) = (\bigcup_{i \in I} U_i) \cap Y$. ▮

Definition. The topology $\tau|_Y$ is called the subspace topology or the topology induced on Y by τ.

Definition. A subset Y of a topological space (X, τ) is said to be closed if $X \setminus Y$ is open.

Warning. Most of the subsets of $\mathbb{R}^n$ are neither open nor closed.

Definition. We say that a subset U of X is a neighborhood of a point $x \in X$, if there is an open subset V such that $x \in V \subset U$.

Definition. Let (X,τ) and (Y,τ') be topological spaces. A mapping $f: X \longrightarrow Y$ is said to be continuous at the point $x \in X$ (with respect to the given topologies) if for each neighborhood U of $f(x)$ $f^{-1}(U)$ is a neighborhood of x. The mapping f is continuous (with respect to the given topologies), if it is continuous at each point or equivalently if for each $U \in \tau'$ we have $f^{-1}(U) \in \tau$.

Exercise. Show that for $\mathbb{R}^n$ the above definition is equivalent to the "ε-δ" definition of continuity at a point: $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is continuous at $\mathbf{x} \in \mathbb{R}^n$ with respect to the standard topologies if and only if for any $\varepsilon > 0$ one can find a positive δ such that $|\mathbf{x}-\mathbf{x}'| < \delta$ implies $|f(\mathbf{x})-f(\mathbf{x}')| < \varepsilon$.

The map f is a homeomorphism, if it is a bijection (= one-to-one and onto) such that both f and f^{-1} are continuous.

We say that two topological spaces are homeomorphic or have the same topological type, if there is a homeomorphism between them.

Homeomorphic topological spaces are considered to be the same from the viewpoint of topology. Intuitively, two homeomorphic spaces are homeomorphic if a rubber model of one of them can be transformed into that of the other. We are allowed to stretch and shrink the model but not allowed to cut the model or glue pieces together. More exactly, we may cut the model somewhere only if later on we glue together the parts we get in the same way as they were joined. Of course, this description of homeomorphism is applicable only for "nice spaces" such as surfaces, curves etc. and by no means substitutes the precise definition.

For example, the circle, the perimeter of a square, and the trefoil knot are homeomorphic, so are a solid disc and a solid square, however a circle is not homeomorphic to a solid disc. Generally it is easy to show that two homeomorphic spaces are indeed homeomorphic: we only have to present a homeomorphism. However, to show that two topological spaces are not homeomorphic, we have to find a topological property, which is had by one of the spaces but is not by the other. For example, the fact that $\mathbb{R}^n$ is not homeomorphic to $\mathbb{R}^m$ if $n \neq m$, is a deep theorem of topology (the "dimension invariance theorem"), the proof of which requires sophisticated techniques.

CURVES IN $\mathbb{R}^n$

Definition. A simple arc in a topological space is a subset Γ homeomorphic to a closed interval $[a,b]$ of $\mathbb{R}$. A parameterization of a simple arc is a homeomorphism $\gamma : [a,b] \longrightarrow \Gamma$.

Definition. A path or a continuous (parameterized) curve in a topological space is a continuous mapping of an interval $[a,b]$ into the space. The images of a and b are the initial and terminal points of the curve respectively. The path is said to connect the initial point to the terminal point.

As we see, there are two different approaches to the concept of curve. According to the intuitive-geometrical approach, curves are *topological spaces* or *subsets* of a topological space. The second approach however, which will be more suitable for our purposes, introduces curves as trajectories of a moving point. This view is reflected in the definition of a continuous curve. We stress that according to this definition, a curve is a *mapping* and not a set of points.

Remark. The image of a continuous curve $\mathbf{x} : [a,b] \longrightarrow \mathbb{R}^n$ as a point set may not look like a curve. The Italian mathematician Peano (1858-1932) constructed a continuous curve that passes through each point of a square. Such a pathology can not occur if we restrict ourselves to smooth curves.

Definition. A smooth (parameterized) curve in $\mathbb{R}^n$ is a smooth map $\mathbf{x} : [a,b] \longrightarrow \mathbb{R}^n$ from a closed interval $[a,b]$ into $\mathbb{R}^n$.

Recall that a map $\mathbf{x} : U \longrightarrow \mathbb{R}^n$ defined on an open subset of $\mathbb{R}^n$ is said to be smooth or infinitely many times differentiable if the coordinate functions $x_1, x_2, \dots, x_n$ of $\mathbf{x} = (x_1, x_2, \dots x_n)$ have continuous partial derivatives of any order.

We say that a map $\mathbf{x} : A \longrightarrow \mathbb{R}^n$ defined on an *arbitrary* set $A \subset \mathbb{R}^n$ is smooth or infinitely many times differentiable if there exists an open set $U \subset \mathbb{R}^n$ and a smooth mapping $\tilde{\mathbf{x}} : U \longrightarrow \mathbb{R}^n$ such that $A \subset U$ and $\mathbf{x} = \tilde{\mathbf{x}}|_A$.

Definition. We say that the continuous curve $\mathbf{x}_1 : [a,b] \longrightarrow \mathbb{R}^n$ is obtained from the curve $\mathbf{x}_2 : [c,d] \longrightarrow \mathbb{R}^n$ by a reparameterization if there is a homeomorphism $\varphi : [a,b] \longrightarrow [c,d]$ such that $\varphi(a) = c$, $\varphi(b) = d$ and $\mathbf{x}_1 = \mathbf{x}_2 \circ \varphi$. A reparameterization is called regular if φ is smooth and $\varphi' > 0$.

Intuitively, reparameterization means that without changing the trajectory of a point we change the velocity, with which it moves along the trajectory.

Definition. The length of a continuous curve $\mathbf{x} : [a,b] \longrightarrow \mathbb{R}^n$ is the limit of the length of inscribed broken lines with vertices $\mathbf{x}(t_0), \mathbf{x}(t_1), \dots, \mathbf{x}(t_N)$ where $a = t_0 < t_1 < \dots < t_N = b$ and the limit is taken as $\max_{1 \leqslant i \leqslant N} |t_i - t_{i-1}|$ tends to zero. Provided that this limit exists (it does not have to), the curve is called rectifiable.

The following theorem yields a formula that can be used in practice to compute the length of curves.

Theorem. A smooth curve $\mathbf{x} : [a,b] \longrightarrow \mathbb{R}^n$ is always rectifiable and its length is determined by the integral

$$\ell(\mathbf{x}) = \int_a^b \| \mathbf{x}'(t) \| \, dt.$$

Proof. Denoting by $x_1, x_2, \dots, x_n$ the coordinate functions of $\mathbf{x}$ the length of the broken line considered in the definition of the length of a curve is equal to

$$\lambda = \sum_{i=1}^{N} \sqrt{\sum_{j=1}^{n} (x_j(t_i) - x_j(t_{i-1}))^2}$$

By Lagrange's mean value theorem we can find real numbers ξ_{ij} such that

$$x_j(t_i) - x_j(t_{i-1}) = x_j'(\xi_{ij})\,(t_i - t_{i-1}), \quad t_{i-1} < \xi_{ij} < t_i.$$

Using these equalities we get

$$\lambda = \sum_{i=1}^{N} \left(\sqrt{\sum_{j=1}^{n} x_j'(\xi_{ij})^2} \right) (t_i - t_{i-1}).$$

Lemma. A continuous function $f : [a,b] \longrightarrow \mathbb{R}$ on a closed interval is uniformly continuous (i.e. for each positive ε one can find a positive δ such that $x, y \in [a,b]$, $|x-y| < \delta$ implies

$$|f(x) - f(y)| < \varepsilon.$$

Proof of the lemma. Suppose to the contrary, that there is an $\varepsilon > 0$ for which we can not find a suitable δ. Then there exists a sequence of pairs of real numbers x_n, y_n such that $x_n, y_n \in [a,b]$, $|x_n - y_n| < 1/n$ but $|f(x_n)-f(y_n)| > \varepsilon$. By compactness of $[a,b]$, we can select a convergent subsequence $x_{i_n} \longrightarrow x$ of the sequence (x_n). Condition $|x_n - y_n| < 1/n$ ensures that $y_{i_n} \longrightarrow x$ as well and so, by the continuity of f at x we have $|f(x_{i_n})-f(y_{i_n})| \longrightarrow 0$. But this contradicts the condition $|f(x_n)-f(y_n)|>\varepsilon$. ▌

Let us return to the proof of the theorem. Fix a positive ε. By the lemma, we can find a positive δ such that $t,t' \in [a,b]$ and $|t - t'| < \delta$ imply $|x_j'(t) - x_j'(t')| < \varepsilon$ for all j.

Suppose that the approximating broken line is fine enough in the sense that $|t_i - t_{i-1}| < \delta$ for all i. Then we have by the triangle inequality

$$\left| \sqrt{\sum_{j=1}^{n} x_j'(\xi_{ij})^2} - \sqrt{\sum_{j=1}^{n} x_j'(t_i)^2} \right| \leqslant$$

$$\leqslant \sqrt{\sum_{j=1}^{n} (x_j'(\xi_{ij})-x_j'(t_i))^2} \leqslant \varepsilon\sqrt{n}.$$

Making use of this estimation we see that

$$\left| \lambda - \sum_{i=1}^{N} \sqrt{\sum_{j=1}^{n} x_j'(t_i)^2}\,(t_i-t_{i-1}) \right| \leqslant \varepsilon\sqrt{n}\,(b-a) \qquad (*)$$

In this formula

$$\sum_{i=1}^{N} \sqrt{\sum_{j=1}^{n} x_j'(t_i)^2}\,(t_i-t_{i-1}) = \sum_{i=1}^{N} \|\mathbf{x}'(t_i)\|\,(t_i-t_{i-1})$$

is just an integral sum which converges to the integral $\int_a^b \|\mathbf{x}'(t)\|\,dt$ when $\max_i |t_i - t_{i-1}|$ tends to zero. In this case the inequality (*) guarantees that the length λ of the inscribed broken lines also tends to this integral. ▌

Let us consider the function $s\colon [a,b] \longrightarrow [0,\ell(\mathbf{x})]$

$$s(t) = \int_a^t \|\mathbf{x}'(\tau)\|\,d\tau ,$$

where $\mathbf{x}\colon [a,b] \longrightarrow \mathbb{R}^n$ is a smooth curve. $s(t)$ is the length of the arc of the curve between $\mathbf{x}(a)$ to $\mathbf{x}(t)$ and is not a strictly monotonous function of t in general. This fact motivates the following definition.

Definition. A smooth curve is said to be regular if $\mathbf{x}'(t) \neq 0$ for all t.

If $\mathbf{x}$ is a regular curve then s defines a regular reparameterization of $\mathbf{x}$. The map $\mathbf{x}\circ s^{-1}\colon [0,\ell(\mathbf{x})] \longrightarrow \mathbb{R}^n$ is referred to as the natural or unit speed parameterization of the curve $\mathbf{x}$ or as the parameterization of $\mathbf{x}$ by arc length. The second name is justified by the fact that the speed vector

$$(\mathbf{x}\circ s^{-1})'(t) = \mathbf{x}'(s^{-1}(t))\,(s^{-1})'(t) = \mathbf{x}'(s^{-1}(t))\frac{1}{s'(s^{-1}(t))} = \frac{\mathbf{x}'(s^{-1}(t))}{\|\mathbf{x}'(s^{-1}(t))\|}$$

of this parameterization has unit module at each point.

Further Exercises

1-1. Show that $\mathbb{R}^n$ and the open balls in $\mathbb{R}^n$ (with the subspace topology) are homeomorphic.

1-2. Show that the "punctured sphere" $S^2-\{p\}$, where

$$S^2 = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| = 1\} \quad \text{and } p \in S^2$$

furnished with the subspace topology is homeomorphic to the plane $\mathbb{R}^2$.

1-3. Using the intuitive description of homeomorphism classify (without proof) the capital letters of the alphabet up to homeomorphism. (The answer depends on the font type you choose, so choose your favorite font.)

1-4. The curve cycloid is the trajectory of a peripheral point of a circle that rolls along a straight line. Find a parameterization of the cycloid and compute the length of one of its arcs.

1-5. Find the natural parameterization of the helix

$$\gamma(t) = (a \cos t,\ a \sin t,\ b\, t).$$

UNIT 2. Curvatures of a Curve

Convergence of k-planes, the osculating k-plane, curves of general type in $\mathbb{R}^n$, the osculating flag, vector fields, moving frames and Frenet frames along a curve, orientation of a vector space, the standard orientation of $\mathbb{R}^n$, the distinguished Frenet frame, Gram-Schmidt orthogonalization process, Frenet formulas, curvatures, invariance theorems, curves with prescribed curvatures.

One of the most important tools of analysis is linearization, or more generally, the approximation of general objects with easily treatable ones. E.g. the derivative of a function is the best linear approximation, Taylor's polynomials are the best polynomial approximations of the function around a point. Adapting this idea to the theory of curves, the following questions arise naturally. Given a curve $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ and a point $\gamma(\bar{t})$ on it, find the straight line or circle (or conic or polynomial curve of degree $\leqslant n$) that approximates the curve around $\gamma(\bar{t})$ best or find the k-plane (k-sphere, quadric surface etc.) that is tangent to the curve at $\gamma(\bar{t})$ with the highest possible order.

We shall deal now with the problem of finding the k-plane that fits to a curve at a given point best. The classical approach to this problem is the following. A k-plane is determined uniquely by (k+1) of its points that do not lie in a (k-1)-plane. Let us take k+1 points $\gamma(t_0),\gamma(t_1),\dots,\gamma(t_k)$ on the curve. If γ is a curve of "general type" then these points span a unique k-plane, which will be denoted by $A(t_0,\dots,t_k)$. The k-plane we look for is the limit position of the k-planes $A(t_0,\dots,t_k)$ as $t_0,\dots,t_k$ tend to $\bar{t}$. To properly understand the last sentence, we need a definition of convergence of k-planes.

Definition. Let $X_1,X_2,\dots;$ X be k-planes. We say that the sequence of k-planes $X_1,X_2,\dots$ tends to X if one can find points $\mathbf{p}_j \in X_j$ $\mathbf{p} \in X$ and linearly independent direction vectors $\mathbf{v}_1^j,\dots,\mathbf{v}_k^j$ of X_j and $\mathbf{v}_1,\dots,\mathbf{v}_k$ of X such that $\lim \mathbf{p}_j = \mathbf{p}$ and $\lim \mathbf{v}_i^j = \mathbf{v}_i$ for $i = 1,\dots,k$ as j tends to infinity.

Exercise*. Show that a sequence of k-planes can have at most one limit.

Solution. Suppose that the sequence $X_1,X_2,\dots$ has two limits X and Y. Then by the definition, one can find points $\mathbf{p}_j$, $\mathbf{q}_j \in X_j$ $\mathbf{p} \in X$ $\mathbf{q} \in Y$ and linearly independent direction vectors $\{\mathbf{v}_1^j,\dots,\mathbf{v}_k^j\}$ and $\{\mathbf{w}_1^j,\dots,\mathbf{w}_k^j\}$ of X_j and $\{\mathbf{v}_1,\dots,\mathbf{v}_k\}$ of X and $\{\mathbf{w}_1,\dots,\mathbf{w}_k\}$ of Y, such that $\lim \mathbf{p}_j = \mathbf{p}$, $\lim \mathbf{q}_j = \mathbf{q}$,

$\lim \mathbf{v}_i^j = \mathbf{v}_i$ and $\lim \mathbf{w}_i^j = \mathbf{w}_i$ for i = 1,...,k as j tends to infinity. Since $\{\mathbf{v}_1^j,\dots,\mathbf{v}_k^j\}$ and $\{\mathbf{w}_1^j,\dots,\mathbf{w}_k^j\}$ span the same linear space, which contains the direction vector $\mathbf{p}_j-\mathbf{q}_j$, there exists a unique kxk matrix (a_{rs}^j) $1\leqslant r,s\leqslant k$ and a vector $(b_1^j,\dots,b_k^j)$ such that

$$\mathbf{v}_i^j = \sum_{s=1}^{k} a_{is}^j \mathbf{w}_s^j \quad \text{for } i=1,\dots,k. \qquad (*)$$

$$\mathbf{p}_j-\mathbf{q}_j = \sum_{s=1}^{k} b_s^j \mathbf{w}_s^j$$

The components a_{rs}^j of this matrix and the numbers b_s^j can be determined by solving the system (*) of linear equations, thus by Cramer's rule, they are rational functions (quotients of polynomials) of the components of the vectors $\mathbf{v}_i^j$ and $\mathbf{w}_i^j$. The denominator of the quotient expressing a_{rs}^j and b_r^j is a non-vanishing kxk minor of the matrix with rows $\mathbf{w}_1^j,\mathbf{w}_2^j,\dots,\mathbf{w}_k^j$. Using the facts that rational functions are continuous and that the denominator of the fraction expressing a_{rs}^j or b_r^j tends to a non-vanishing minor of the matrix with rows $\mathbf{w}_1,\mathbf{w}_2,\dots,\mathbf{w}_k$ if we choose the minors mentioned above properly, we get that the limits $\lim a_{rs}^j = a_{rs}$, $\lim b_r^j = b_r$ exist as $j\to\infty$. Taking $j\to\infty$ in (*) we obtain

$$\mathbf{v}_i = \sum_{s=1}^{k} a_{is} \mathbf{w}_s \quad \text{for } i=1,\dots,k,$$

$$\mathbf{p}-\mathbf{q} = \sum_{s=1}^{k} b_s \mathbf{w}_s,$$

from which follows that the vectors $\mathbf{v}_1,\dots,\mathbf{v}_k$ and $\mathbf{w}_1,\dots,\mathbf{w}_k$ span the same linear space, i.e. the k-planes X and Y must be parallel and that the point $\mathbf{q}$ is a common point of them consequently X = Y. ▌

<u>Remark</u>. The set of all k-dimensional linear/affine subspaces of an n-dimensional linear space has a natural topology, which can easily be described with the help of the factor space topology construction.

Assume that a topological space (X,τ) is divided into a disjoint union of its subsets. Such a subdivision can always be thought of as a splitting of X into the equivalence classes of an equivalence relation ~ on X. Denoting by $Y = X/_\sim$ the set of equivalence classes we have a natural mapping $\pi: X\longrightarrow Y$ assigning to an element $x \in X$ its equivalence class $[x] \in Y$.

<u>Proposition</u>. The set

$$\tau' = \{ U \subset Y : \pi^{-1}(U) \in \tau \}$$

is a topology on Y.

<u>Proof</u>. The proof follows from the following set theoretical identities.

i) $\pi^{-1}(\emptyset) = \emptyset, \quad \pi^{-1}(Y) = X$;

ii) $\pi^{-1}(U \cap V) = \pi^{-1}(U) \cap \pi^{-1}(V)$;

iii) $\pi^{-1}(\bigcup_{i\in I} U_i) = \bigcup_{i\in I} \pi^{-1}(U_i)$. ▮

<u>Definition</u>. The family τ' is called the <u>factor space topology</u> on Y.

Now consider the set

$$V(n,k) = \{(\mathbf{x}_1,\dots,\mathbf{x}_k) \in (\mathbb{R}^n)^k : \mathbf{x}_1,\dots,\mathbf{x}_k \text{ are linearly independent}\}.$$

$V(n,k)$ is an open subset in $(\mathbb{R}^n)^k = \mathbb{R}^{nk}$ hence it inherits a subspace topology from the standard topology of $\mathbb{R}^{nk}$. If we define two elements of $V(n,k)$ to be equivalent if they span the same k-dimensional linear subspace of $\mathbb{R}^n$, then the set $Gr(n,k) = V(n,k)/_\sim$ of equivalence classes is essentially the same as the set of all k-dimensional linear subspaces of $\mathbb{R}^n$. This set becomes a topological space with the factor space topology. Topological spaces of the form $Gr(n,k)$ are called <u>Grassmann manifolds</u>.

<u>Exercise</u>. Show that $Gr(n,k)$ is homeomorphic to $Gr(n,n-k)$.

We can define <u>affine Grassmann manifolds</u> similarly. We set

$$\tilde{V}(n,k) = \{(\mathbf{x}_0,\dots,\mathbf{x}_k) \in (\mathbb{R}^n)^{k+1} : \mathbf{x}_0,\dots,\mathbf{x}_k \text{ are not in a (k-1)-plane}\},$$

furnish $\tilde{V}(n,k)$ with the subspace topology inherited from $\mathbb{R}^{n(k+1)}$ and define an equivalence relation on $\tilde{V}(n,k)$ by

$$(\mathbf{x}_0,\dots,\mathbf{x}_k) \sim (\mathbf{y}_0,\dots,\mathbf{y}_k) \Leftrightarrow \mathbf{x}_0,\dots,\mathbf{x}_k \text{ and } \mathbf{y}_0,\dots,\mathbf{y}_k \text{ span the same k-plane.}$$

$\tilde{V}(n,k)/_\sim$ is the set of affine k-dimensional subspaces that has the factor space topology on it.

<u>Exercise</u>. Show that convergence of k-planes as defined above is the same as convergence with respect to the topology we have just constructed.

<u>Definition</u>. Let $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ be a curve, $\bar{t} \in (a,b)$, $1 \leqslant k \leqslant n$. If the k-planes $A(t_0,\dots,t_k)$ are defined for parameters close enough to $\bar{t}$ and their limit exists as $t_0,\dots,t_k \to \bar{t}$, then the limit is called the <u>osculating k-plane</u> of the curve γ at $\bar{t}$. The osculating 1-plane of a curve is just the <u>tangent</u> of the curve. (The word "osculate" comes from the Latin "osculari - to kiss").

The definition of the osculating k-plane is justified by intuition. It is the k-plane that passes through k+1 points "infinitely close" to a given point. However, it would not be easy to determine the osculating k-plane using directly its definition. Fortunately, we have the following theorem.

Theorem. Let $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ be a smooth curve, $\bar{t} \in (a,b)$, $1 \leqslant k \leqslant n$. If the derivatives $\gamma'(\bar{t}), \gamma''(\bar{t}), \ldots, \gamma^{(k)}(\bar{t})$ are linearly independent, then the osculating k-plane of γ is defined at $\bar{t}$ and it is the k-plane that passes through $\gamma(\bar{t})$ with direction vectors $\gamma'(\bar{t}), \gamma''(\bar{t}), \ldots, \gamma^{(k)}(\bar{t})$.

To prove the theorem we need some preparation.

Definition. Let $f : [a,b] \longrightarrow \mathbb{R}^n$ be a vector valued function, $t_0, t_1, \ldots \in [a,b]$ are different numbers. The higher order divided differences or difference quotients are defined recursively

$$f_0(t_0) := f(t_0)$$

$$f_1(t_0, t_1) := \frac{f(t_1) - f(t_0)}{t_1 - t_0}$$

$$\ldots$$

$$f_k(t_0, t_1, \ldots, t_k) := \frac{f_{k-1}(t_1, \ldots, t_k) - f_{k-1}(t_0, \ldots, t_{k-1})}{t_k - t_0}$$

Exercise. Show that the k-th order divided difference is a symmetric function of the variables $t_0, \ldots, t_k$ and has the following explicit form

$$f_k(t_0, t_1, \ldots, t_k) = \sum_{i=0}^{k} f(t_i) \frac{1}{\omega'(t_i)} ,$$

where $\omega(t)$ is the polynomial $(t-t_0)(t-t_1)\ldots(t-t_k)$, hence

$$\omega'(t_i) = (t_i-t_0)(t_i-t_1)\ldots(t_i-t_{i-1})(t_i-t_{i+1})\ldots(t_i-t_k).$$

Lemma. If $f : [a,b] \longrightarrow \mathbb{R}$ is a smooth function, then there exists a number $\xi \in [a,b]$ such that

$$f_k(t_0, t_1, \ldots, t_k) = \frac{f^{(k)}(\xi)}{k!} .$$

Proof. Let $P(t)$ be the polynomial of degree $\leqslant k$ for which $f(t_i) = P(t_i)$ for $i=0,1,\ldots,k$. Such a polynomial exists and is unique. P is unique, since if Q is also a polynomial of degree $\leqslant k$ such that $f(t_i) = P(t_i) = Q(t_i)$ for $i=0,\ldots,k$, then the polynomial $P-Q$ has degree $\leqslant k$ and $k+1$ roots, which is possible only in the case when $P-Q \equiv 0$, $P \equiv Q$. We show the existence of P by an explicit construction. Set

$$P_i(t) = \frac{(t-t_0)(t-t_1)\ldots(t-t_{i-1})(t-t_{i+1})\ldots(t-t_k)}{(t_i-t_0)(t_i-t_1)\ldots(t_i-t_{i-1})(t_i-t_{i+1})\ldots(t_i-t_k)} .$$

Obviously, P_i is a polynomial of degree k such that $P_i(t_j) = \delta_{ij}$ (Kronecker δ symbol denotes $\delta_{ij} = 1$ if $i=j$ and $\delta_{ij} = 0$ if $i \neq j$). Thus, the polynomial

$$P(t) = \sum_{i=0}^{k} f(t_i) P_i(t)$$

is good for our purposes.

The difference f−P has k+1 roots. Since by the mean value theorem the interval between two zeros of a smooth function contains an interior point at which the derivative vanishes, the derivative f'− P' has at least k roots. Similarly, f''−P'' has at least k−1 roots, etc. $f^{(k)}-P^{(k)}$ vanishes at a certain point $\xi \in [a,b]$. The k-th derivative of a polynomial of degree k is k! times the coefficient of the highest power, from which

$$f^{(k)}(\xi) = P^{(k)}(\xi) = k! \sum_{i=0}^{k} f(t_i) \frac{1}{\omega'(t_i)} = k!\ f_k(t_0,t_1,\dots,t_k). \blacksquare$$

Corollary. Since ξ can be chosen from the interval spanned by the points $t_0,\dots,t_k$, if these points tend to $\bar{t} \in [a,b]$, then ξ also tends to $\bar{t}$, consequently $f_k(t_0,t_1,\dots,t_k) = \frac{f^{(k)}(\xi)}{k!}$ tends to $\frac{f^{(k)}(\bar{t})}{k!}$.

Corollary. Applying the previous corollary to the components of a vector valued function $f:[a,b]\longrightarrow \mathbb{R}^n$, we get $f_k(t_0,t_1,\dots,t_k)$ tends to $\frac{f^{(k)}(\bar{t})}{k!}$ as $t_0,\dots,t_k$ tend to $\bar{t}$.

Proof of the theorem. Let us recall that if $\mathbf{p}_0,\dots,\mathbf{p}_k$ are position vectors of k+1 points in $\mathbb{R}^n$, then the affine plane spanned by them consists of linear combinations the coefficients in which have sum equal to 1

$$A(\mathbf{p}_0,\dots,\mathbf{p}_k) = \{\alpha_0\mathbf{p}_0+\dots+\alpha_k\mathbf{p}_k \ :\ \alpha_0+\dots+\alpha_k=1\ \}.$$

The direction vectors of this affine plane are linear combinations $\alpha_0\mathbf{p}_0+\dots+\alpha_k\mathbf{p}_k$ such that $\alpha_0+\dots+\alpha_k=0$.

Exercise. Prove this.

We claim that if $\gamma:[a,b]\longrightarrow \mathbb{R}^n$ is a curve in $\mathbb{R}^n$, then $\gamma_k(t_0,\dots,t_k)$ is a direction vector of the affine linear subspace spanned by the points $\gamma(t_0),\dots\gamma(t_k)$. To see this, we have to show that $\sum_{i=0}^{k} \frac{1}{\omega'(t_i)} = 0$. Consider the function $f\equiv 1$ and construct the polynomial P of degree $\leqslant k$ which coincides with f at $t_0,\dots,t_k$ using the general formulae. By the above proposition, there exists a number ξ such that

$$0 = f^{(k)}(\xi) = P^{(k)}(\xi) = \sum_{i=0}^{k} \frac{1}{\omega'(t_i)} ,$$

as we wanted to show. This way, $\gamma(t_0)$ is a point and $\gamma_1(t_1,t_0),\dots,\gamma_k(t_k,\dots,t_0)$ are direction vectors of the affine linear subspace spanned by the points $\gamma(t_0),\dots,\gamma(t_k)$. If γ is smooth, then $\gamma(t_0)$ tends to $\gamma(\bar{t})$, $\gamma_1(t_1,t_0)$ tends to $\gamma'(\bar{t})$, and so on, $\gamma_k(t_k,\dots,t_0)$ tends to $\gamma^{(k)}(\bar{t})/k!$ as the points $t_0,\dots,t_k$ tend to $\bar{t} \in [a,b]$. Since by our assumption the first k derivatives of γ are linearly independent at $\bar{t}$, so are the vectors $\gamma_1(t_1,t_0),\dots,\gamma_k(t_k,\dots,t_0)$ if $t_0,\dots,t_k$ are in a small neighborhood of $\bar{t}$ and in this case the k-plane $A(t_0,\dots,t_k)$ tends to the k-plane that passes through $\gamma(\bar{t})$ with direction vectors $\gamma'(\bar{t}),\gamma''(\bar{t})/2,\dots,\gamma^{(k)}(\bar{t})/k!$. ▮

Definition. Let V be an n-dimensional vector space. A flag in V is a sequence of linear subspaces $\{\mathbf{0}\}=V_0\subset V_1\subset\dots\subset V_n=V$ such that $\dim V_i = i$. An affine flag is a sequence of affine subspaces $A_0\subset A_1\subset\dots\subset A_n=V$ such that $\dim A_i = i$.

Definition. A curve $\gamma:[a,b]\longrightarrow \mathbb{R}^n$ is called a curve of general type in $\mathbb{R}^n$ if the first n-1 derivatives $\gamma'(t),\gamma''(t),\dots,\gamma^{(n-1)}(t)$ are linearly independent for all $t \in [a,b]$.

Definition. The osculating flag of a curve of general type at a given point is the affine flag consisting of the osculating k-planes for $k = 0,1,\dots,n-1$ and the whole space.

Our plan is the following. A curve of general type in $\mathbb{R}^n$ is not contained in any affine subspace of dimension $k < n-1$, so we may pose the question, how far is it from being contained in a k-plane. In other words, we want to measure the deviation of the curve from its osculating k-plane. One way to do this is that we measure how quickly the osculating flag rotates as we travel along the curve. Since the faster we travel along the curve the faster change we observe, it is natural to consider the speed of rotation of the osculating flag with respect to the unit speed parameterization of the curve. This will lead us to quantities that describe the way a curve is curved in space, which are called the curvatures of the curve. There is one question of technical character left: how can we measure the change of an affine subspace. This problem can be solved by introducing an orthonormal basis at each point in such a way that the first k basis vectors span the osculating k-plane at the point in question, then measuring the change of this basis.

Definition. A (smooth) vector field along a curve $\gamma:[a,b]\longrightarrow \mathbb{R}^n$ is a smooth mapping $\mathbf{v}:[a,b]\longrightarrow T_*\mathbb{R}^n$ such that $\mathbf{v}(t) \in T_{\gamma(t)}\mathbb{R}^n$ for all $t \in [a,b]$. (We shall deal only with smooth vector fields.)

Thus, $\mathbf{v}(t)$ is a vector based at $\gamma(t)$. If we forget about the initial point of $\mathbf{v}(t)$, which is, after all, determined by the parameter t, then $\mathbf{v}$ can be considered as simply a mapping from the parameter interval [a,b] to $\mathbb{R}^n$.

Definition. A moving frame along a curve $\gamma:[a,b]\longrightarrow \mathbb{R}^n$ is a collection of n vector fields $\mathbf{t}_1,\dots,\mathbf{t}_n$ along γ such that $\langle\mathbf{t}_i(t),\mathbf{t}_j(t)\rangle = \delta_{ij}$ for all $t \in [a,b]$.

There are many moving frames along a curve and most of them have nothing to do with the geometry of the curve. This is not the case for Frenet frames.

Definition. A moving frame $\mathbf{t}_1,\ldots,\mathbf{t}_n$ along a curve γ is called a Frenet frame if for all k, $1\leqslant k\leqslant n$, $\gamma^{(k)}(t)$ is contained in the linear span of $\mathbf{t}_1(t),\ldots,\mathbf{t}_k(t)$.

Exercise. Construct a curve which has no Frenet frame and one with infinitely many Frenet frames. Show that a curve of general type in $\mathbb{R}^n$ has exactly 2^n Frenet frames.

According to the exercise, a Frenet frame along a curve of general type is almost unique. To select a distinguished Frenet frame from among all of them, we use orientation.

Definition. Let $\mathbf{v}_1,\ldots,\mathbf{v}_k$ and $\mathbf{w}_1,\ldots,\mathbf{w}_k$ be two ordered bases of a linear space V. We say that they have the same orientation or they define the same orientation of V, if the kxk matrix (a_{ij}) defined by the system of equalities

$$\mathbf{v}_i = \sum_{j=1}^{k} a_{ij}\,\mathbf{w}_j \qquad \text{for } i = 1,2,\ldots,k$$

has positive determinant.

Having the same orientation is an equivalence relation on ordered bases, and there are two equivalence classes. Fixing one of the equivalence classes the elements of which will be called then positively oriented bases is an orientation of V.

Definition. The standard orientation of $\mathbb{R}^n$ is the orientation defined by the ordered basis $\mathbf{e}_1,\ldots,\mathbf{e}_n$, where $\mathbf{e}_i = (0,..,0,\underset{\text{i-th place}}{\hat{1}},0,\ldots,0)$.

Definition. A Frenet frame $\mathbf{t}_1,\ldots,\mathbf{t}_n$ of a curve γ of general type in $\mathbb{R}^n$ is called a distinguished Frenet frame if for all k, $1\leqslant k\leqslant n-1$, the vectors $\mathbf{t}_1(t),\ldots,\mathbf{t}_k(t)$ have the same orientation in their linear span as the vectors $\gamma'(t),\ldots,\gamma^{(k)}(t)$, and the basis $\mathbf{t}_1(t),\ldots,\mathbf{t}_n(t)$ is positively oriented with respect to the standard orientation of $\mathbb{R}^n$.

Proposition. A curve of general type possesses a unique distinguished Frenet frame.

Proof. We can determine the first n-1 vector fields of the distinguished Frenet frame by application of the Gram-Schmidt orthogonalization process to the first n-1 derivatives of γ. According to this, we set

$$\mathbf{t}_1 = \frac{\gamma'}{\|\gamma'\|}.$$

Suppose that $\mathbf{t}_1,\ldots,\mathbf{t}_k$ have already been defined. Then $\mathbf{t}_{k+1}$ must be of the form

$$\mathbf{t}_{k+1} = \beta\gamma^{(k+1)} + (\alpha_1\mathbf{t}_1+\ldots+\alpha_k\mathbf{t}_k), \qquad (*)$$

where the coefficients β, $\alpha_1,\dots,\alpha_k$ are to be determined. Taking the dot product of both sides of (*) with $\mathbf{t}_i$ $(1\leqslant i\leqslant k)$, we obtain

$$0 = \beta <\gamma^{(k+1)},\mathbf{t}_i> + \alpha_i,$$

consequently,

$$\mathbf{t}_{k+1} = \beta \left(\gamma^{(k+1)} -(<\gamma^{(k+1)},\mathbf{t}_1>\mathbf{t}_1+\dots+<\gamma^{(k+1)},\mathbf{t}_k>\mathbf{t}_k)\right).$$

The parameter β must be used to normalize the vector which stands on the right of it. Thus,

$$\beta = \pm \| \gamma^{(k+1)} -(<\gamma^{(k+1)},\mathbf{t}_1>\mathbf{t}_1+\dots+<\gamma^{(k+1)},\mathbf{t}_k>\mathbf{t}_k)\|^{-1}.$$

Exercise. Show that in order to get a distinguished Frenet frame, we have to choose a positive β, i.e.

$$\mathbf{t}_{k+1} = \frac{\left(\gamma^{(k+1)} -(<\gamma^{(k+1)},\mathbf{t}_1>\mathbf{t}_1+\dots+<\gamma^{(k+1)},\mathbf{t}_k>\mathbf{t}_k)\right)}{\|\gamma^{(k+1)} -(<\gamma^{(k+1)},\mathbf{t}_1>\mathbf{t}_1+\dots+<\gamma^{(k+1)},\mathbf{t}_k>\mathbf{t}_k)\|}.$$

To finish the proof, we have to show that given n-1 mutually orthogonal unit vectors $\mathbf{t}_1,\dots,\mathbf{t}_{n-1}$ in $\mathbb{R}^n$, there is a unique vector $\mathbf{t}_n$ for which the vectors $\mathbf{t}_1,\dots,\mathbf{t}_n$ form a positively oriented orthonormal basis of $\mathbb{R}^n$. The condition that a vector is perpendicular to $\mathbf{t}_1,\dots,\mathbf{t}_{n-1}$ is equivalent to a system of n-1 linearly independent linear equation, the solutions of which form a 1-dimensional linear subspace (a straight line). There are exactly two opposite unit vectors parallel to a given straight line, and exactly one of them will fulfill the orientation condition. (Replacing a vector of an ordered basis by its opposite changes the orientation.) ▮

Exercise. Show that if $\mathbf{e}_1,\dots,\mathbf{e}_n$ is the standard basis of $\mathbb{R}^n$ and

$$\mathbf{t}_i=\alpha_{i1}\mathbf{e}_1+\dots+\alpha_{in}\mathbf{e}_n \text{ for } i=1,\dots,n-1,$$

then $\mathbf{t}_n$ can be obtained as the formal determinant of the matrix

$$\begin{vmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ & \dots & \\ \alpha_{(n-1)1} & \dots & \alpha_{(n-1)n} \\ \mathbf{e}_1 & & \mathbf{e}_n \end{vmatrix}.$$

Let γ be a curve of general type in $\mathbb{R}^n$ parameterized by arc length. Set

$$\dot{\mathbf{t}}_i = \sum_{j=1}^{n} \alpha_{ij}\ \mathbf{t}_i.$$

(By convention, derivation with respect to arc length is denoted by $\dot{}$ instead of ' .)

Proposition. Using the above notation, $\alpha_{ij} = 0$ provided $j > i+1$.

Proof. Since by construction the vector $\mathbf{t}_i$, $1\leqslant i\leqslant n-1$, is a linear combination of the vectors $\dot\gamma,\dots,\gamma^{(i)}$, $\dot{\mathbf t}_i$ is a linear combination of the vectors $\dot\gamma,\dots,\gamma^{(i+1)}$. Since the last vectors are linearly expressible in terms of the vectors $\dot{\mathbf t},\dots,\mathbf t^{(i+1)}$, this proves the proposition. ∎

Proposition. The matrix (α_{ij}) is skew-symmetric i.e. $\alpha_{ij} = -\alpha_{ji}$.

Proof. Since $\langle \mathbf t_i, \mathbf t_j\rangle \equiv \delta_{ij}$ is a constant function, differentiating we get

$$\langle \dot{\mathbf t}_i, \mathbf t_j\rangle + \langle \mathbf t_i, \dot{\mathbf t}_j\rangle = \alpha_{ij} + \alpha_{ji} = 0. \ \blacksquare$$

Thus the only non-zero coefficients are $\alpha_{i,i+1} = -\alpha_{i+1,i}$. Setting

$$\kappa_1 = \alpha_{12},\ \kappa_2 = \alpha_{23}\ ,\dots,\ \kappa_{n-1} = \alpha_{n-1,n}$$

we therefore see that the following formulas hold

$$\begin{aligned}
\dot{\mathbf t}_1 &= \kappa_1\, \mathbf t_2\\
\dot{\mathbf t}_2 &= -\kappa_1\, \mathbf t_1 + \kappa_2\, \mathbf t_3\\
\dot{\mathbf t}_{n-1} &= -\kappa_{n-2}\, \mathbf t_{n-2} + \kappa_{n-1}\, \mathbf t_n\\
\dot{\mathbf t}_n &= -\kappa_{n-1}\, \mathbf t_{n-1}.
\end{aligned}$$

These formulas are called Frenet formulas for a curve in $\mathbb{R}^n$. The functions $\kappa_1,\dots,\kappa_{n-1}$ are called the curvatures of a curve.

Exercise. Show that $\kappa_1,\dots,\kappa_{n-2}$ are positive, while κ_{n-1} may have any sign.

Frenet formulas can easily be modified for smooth curves parameterized in an arbitrary way. If γ is a curve of general type, then it is regular, hence we may consider its reparameterization by arc length $\tilde\gamma = \gamma\circ s^{-1}$. If $\tilde\kappa_1,\dots,\tilde\kappa_{n-1}$ are the curvature functions of the unit speed curve $\tilde\gamma$ defined as above, then we define the curvature functions of the curve γ to be the functions $\kappa_1 = \tilde\kappa_1\circ s,\dots,\kappa_{n-1} = \tilde\kappa_{n-1}\circ s$.

If $\mathbf x$ is an arbitrary vector field along the curve γ, s denotes the arc length parameter, t denotes an arbitrary parameter, then $\frac{d\mathbf x}{dt} = \frac{d\mathbf x}{ds}\frac{ds}{dt} = \frac{d\mathbf x}{ds}\|\gamma'\|$. Thus, denoting by v the length of the speed vector γ', we obtain the following more general formulas

$$\begin{aligned}
\mathbf t'_1 &= v\, \kappa_1\, \mathbf t_2\\
\mathbf t'_2 &= v\, (-\kappa_1\, \mathbf t_1 + \kappa_2\, \mathbf t_3)\\
&\ \ \cdots\\
\mathbf t'_{n-1} &= v\, (-\kappa_{n-2}\, \mathbf t_{n-2} + \kappa_{n-1}\, \mathbf t_n)\\
\mathbf t'_n &= -v\, \kappa_{n-1}\, \mathbf t_{n-1}.
\end{aligned}$$

We formulate some theorems concerning the curvatures of a curve. The first two theorems are intuitively clear and can be proved mechanically. The third theorem is of great theoretical importance, the proof of which uses the existence and uniqueness theorem for the solution of linear differential equations.

Proposition. (Invariance under isometries) Let γ be a curve of general type in $\mathbb{R}^n$, $I:\mathbb{R}^n \longrightarrow \mathbb{R}^n$ an isometry (distance preserving bijection). Then the curvature functions $\kappa_1,\dots,\kappa_{n-2}$ of the curves γ and $I\circ\gamma$ are the same. The last curvatures κ_{n-1} of these curves coincide if I is orientation preserving and they differ (only) in sign if I is orientation reversing.

Proposition. (Invariance under reparameterization) If $\tilde{\gamma}$ is a regular reparameterization of the curve γ i.e. $\tilde{\gamma} = \gamma\circ h$ for some strictly monotone function h, then the curvature functions of $\tilde{\gamma}$ and γ are related to one another by $\tilde{\kappa}_i = \kappa_i\circ h$.

Exercise.

a) Assume $h:[a,b]\longrightarrow [c,d]$ is a smooth bijection between the intervals [a,b] and [c,d] with $h'<0$, $\gamma:[c,d]\longrightarrow \mathbb{R}^n$ is a curve of general type. How are the curvatures of γ and $\tilde{\gamma} = \gamma\circ h$ related to one another?

b) Assume a curve γ of general type in $\mathbb{R}^n$ lies in $\mathbb{R}^{n-1} \subset \mathbb{R}^n$. Then we can compute the curvatures $\kappa_1,\dots,\kappa_{n-1}$ of this curve considering γ to be a curve in $\mathbb{R}^n$ and also we may compute the curvatures $\tilde{\kappa}_1,\dots,\tilde{\kappa}_{n-2}$ of this curve considering γ to be a curve in $\mathbb{R}^{n-1}$. What is the relationship between these two sets of numbers?

Theorem. Given n-2 positive smooth functions $\kappa_1,\dots,\kappa_{n-2}$ and a smooth function κ_{n-1} on an interval [a,b], there exists a unit speed curve of general type in $\mathbb{R}^n$ the curvatures of which are the prescribed functions $\kappa_1,\dots,\kappa_{n-1}$. This curve is unique up to isometries of the space.

Further Exercises

2-1. Describe the curve, called astroid, given by the parameterization $\gamma:[0,2\pi]\longrightarrow \mathbb{R}^2$, $\gamma(t) = (\cos^3(t), \sin^3(t))$. Is the astroid a smooth curve? Is it regular? Is it a curve of general type? If the answer is no for a property, characterize those arcs of the astroid which have the property. Compute the length of the astroid. Show that the segment of a tangent lying between the axis intercepts has the same length for all tangents.

2-2. Show that curvatures $\kappa_1, \ldots, \kappa_{n-2}$ of a curve of general type in $\mathbb{R}^n$ are positive.

2-3. Find distinguished Frenet's basis and the equation of the osculating 2-plane of the elliptical helix $t \longmapsto (a \cos t,\ b \sin t,\ c\ t)$ at the point $(a,0,0)$ (a, b and c are given positive numbers).

2-4. Suppose that a curve of general type in $\mathbb{R}^n$ is contained in an $n-1$ dimensional affine subspace. Show that $\kappa_{n-1} \equiv 0$.

UNIT 3. PLANE CURVES

Explicit formulas for plane curves, rotation number of a closed curve, osculating circle, evolute, involute, parallel curves, "Umlaufsatz". Convex curves and their characterization, the Four Vertex Theorem.

This unit and the following one are devoted to the study of curves in low dimensional spaces. We start with plane curves.

A plane curve $\gamma : [a,b] \longrightarrow \mathbb{R}^2$ is given by two coordinate functions.

$$\gamma(t) = (\ x(t)\ ,\ y(t)\ ) \qquad t \in [a,b].$$

The curve γ is of general position if the vector γ' is a linearly independent "system of vectors". Since a single vector is linearly independent if and only if it is non-zero, the condition of being a curve of general type is equivalent to regularity for plane curves. From this point on we suppose that γ is regular.

The Frenet vector fields $\mathbf{t}_1, \mathbf{t}_2$ are denoted by $\mathbf{t}$ and $\mathbf{n}$ in classical differential geometry and they are called the (unit) tangent and the (unit) normal vector fields of the curve. There is only one curvature function of a plane curve $\kappa = \kappa_1$. Frenet formulas have the form

$$\mathbf{t}' = v\,\kappa\,\mathbf{n}\ ,$$

$$\mathbf{n}' = -\,v\,\kappa\,\mathbf{t}\ ,$$

where $v = |\gamma\,'|$.

Now let us find explicit formulas for $\mathbf{t}$, $\mathbf{n}$, κ. Obviously,

$$\mathbf{t} = \frac{1}{v}\,(x',y') = \frac{1}{\sqrt{x'^2+y'^2}}\,(x',y').$$

The normal vector $\mathbf{n}$ is the last vector of the Frenet basis so it is determined by the condition that $(\mathbf{t},\mathbf{n})$ is a positively oriented orthonormal basis, that is, in our case, $\mathbf{n}$ is obtained from $\mathbf{t}$ by a right angled rotation of positive direction. The right angled rotation in the positive direction takes the vector (a,b) to the vector (-b,a) (check this!) thus

$$\mathbf{n} = \frac{1}{v}\,(-y',x') = \frac{1}{\sqrt{x'^2+y'^2}}\,(-y',x').$$

To express κ, let us start from the equation

$$\gamma' = v\,\mathbf{t}.$$

Differentiating and using the first Frenet formula,

$$\gamma'' = v'\,\mathbf{t} + v\,\mathbf{t}' = v'\mathbf{t} + v^2\kappa\,\mathbf{n}.$$

Taking dot product with $\mathbf{n}$ and using $\langle\mathbf{t},\mathbf{n}\rangle = 0$ we get

$$\langle\gamma'',\mathbf{n}\rangle = v^2\kappa\ ,$$

which gives

$$\kappa = \frac{\langle\gamma'',\mathbf{n}\rangle}{v^2} = \frac{-x''y' + y''x'}{v^3} = \frac{\det\begin{vmatrix} x' & y' \\ x'' & y'' \end{vmatrix}}{(x'^2+y'^2)^{3/2}}\ .$$

Now we establish some facts concerning the curvature of a plane curve which probably will contribute to a better understanding of its geometrical meaning.

First we shall consider curves in an arbitrary dimensional space and investigate the following question. Given a point on a curve, find the circle which approximates the curve around the point with the highest possible accuracy. We have already solved a similar problem for k-planes and the scheme of our approach works for circles as well.

Definition. Let $C_1,C_2,\dots;C$ be circles in $\mathbb{R}^n$, S_i and S the planes of C_i and C resp., O_i and O the centers of C_i and C resp., r_i and r the radii of C_i and C resp. We say that the sequence $C_1,C_2,\dots$ tends to the circle C if the limits $\lim S_i$, $\lim O_i$, $\lim r_i$ exist and equal to S, O, r respectively.

Definition. Let $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ be a smooth curve, $\bar{t} \in [a,b]$, and for any three different arguments $t_1,t_2,t_3 \in [a,b]$, denote by $C(t_1,t_2,t_3)$ the circle passing through $\gamma(t_1),\gamma(t_2),\gamma(t_3)$, provided that these points do not lie in a straight line. If the circles $C(t_1,t_2,t_3)$ are defined if t_1,t_2,t_3 lie in a sufficiently small neighborhood of $\bar{t}$, and their limit $\lim C(t_1,t_2,t_3) = C$ exists as t_1,t_2,t_3 tend to $\bar{t}$, then C is called the osculating circle of the curve γ at the point $\bar{t}$.

Theorem. Suppose that $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ is a smooth curve of general type, $\bar{t} \in [a,b]$ is such that $\kappa_1(\bar{t}) \neq 0$. Then the osculating circle of γ at $\bar{t}$ exists, its plane is the osculating 2-plane of γ at $\gamma(\bar{t})$, its center is the point $\gamma(\bar{t}) + \dfrac{1}{\kappa_1(\bar{t})}\,\mathbf{t}_2(\bar{t})$, its radius is $\dfrac{1}{|\kappa_1(\bar{t})|}$.

Proof. Let us show first that $\gamma(t_1),\gamma(t_2),\gamma(t_3)$ do not lie in a straight line if t_1,t_2,t_3 are in a small neighborhood of $\bar{t}$. We know that the first and second order differences $\gamma_1(t_1,t_2)$ and $\gamma_2(t_1,t_2,t_3)$ tend to $\gamma'(\bar{t})$ and $\gamma''(\bar{t})/2$ respectively as t_1,t_2,t_3 tend to $\bar{t}$. Since $\gamma'(\bar{t})$ and $\gamma''(\bar{t})/2$ are not parallel, (otherwise γ would not be of general type when $n \geqslant 3$ or the curvature

$\kappa_1(\bar{t})$ would be zero when n = 2), neither are $\gamma_1(t_1,t_2)$ and $\gamma_2(t_1,t_2,t_3)$ if t_1,t_2,t_3 are close enough to $\bar{t}$. In this case the affine subspace spanned by $\gamma(t_1),\gamma(t_2)$ and $\gamma(t_3)$ has two non-parallel direction vectors, $\gamma_1(t_1,t_2)$ and $\gamma_2(t_1,t_2,t_3)$, thus it can not be a straight line.

Now suppose that $t_1<t_2<t_3$ are sufficiently close to $\bar{t}$ to guarantee that the circle $C(t_1,t_2,t_3)$ exists and denote by $\mathbf{p}$ and r its center and radius.

The plane of $C(t_1,t_2,t_3)$ is the plane spanned by $\gamma(t_1),\gamma(t_2),\gamma(t_3)$ and this plane converges to the osculating plane at $\bar{t}$.

Set

$$F(t) = |\gamma(t) - \mathbf{p}|^2 - r^2.$$

Since F has three different roots, applying the mean value theorem we find two numbers ξ_1 and ξ_2 such that $t_1<\xi_1<t_2<\xi_2<t_3$ and $F'(\xi_1) = F'(\xi_2) = 0$. By another application of the mean value theorem for F' we find $\xi_1< \eta < \xi_2$ such that $F''(\eta) = 0$. The equations $F'(\xi_1) = F''(\eta) = 0$ give

$$2 <\gamma'(\xi_1),\gamma(\xi_1)-\mathbf{p}> = 0, \qquad (*)$$

and

$$<\gamma''(\eta),\gamma(\eta)-\mathbf{p}> + <\gamma'(\eta),\gamma'(\eta)> = 0. \qquad (**)$$

The first equation means that $\mathbf{p}$ can be written in the form $\mathbf{p} = \gamma(\xi_1) + \rho\ \mathbf{m}$, where $\rho \in \mathbb{R}$, $\mathbf{m}$ is a unit vector perpendicular to $\gamma'(\xi_1)$. Choosing a right orientation for $\mathbf{m}$, we may also assume that $<\mathbf{m},\mathbf{t}_2(\bar{t})>$ is not negative.

<u>Lemma</u>. $\mathbf{m}$ tends to $\mathbf{t}_2(\bar{t})$ as t_1,t_2,t_3 tend to $\bar{t}$.

To prove the lemma, let us write $\mathbf{m}$ as a linear combination of the Frenet basis at $\bar{t}$

$$\mathbf{m} = \sum \alpha_i \mathbf{t}_i(\bar{t}),\ \text{where}\ \alpha_i = <\mathbf{m},\mathbf{t}_i(\bar{t})>.$$

a) $\alpha_1 = <\mathbf{m},\mathbf{t}_1(\bar{t})> = <\mathbf{m},\mathbf{t}_1(\bar{t})-\mathbf{t}_1(\xi_1)>$ tends to 0 since $\mathbf{t}_1(\bar{t})-\mathbf{t}_1(\xi_1)$ tends to $\mathbf{0}$.

b) Let $i \geqslant 3$. Since $\tilde{\mathbf{m}} = \mathbf{m} + (\gamma(\xi_1)-\gamma(t_1))$ is a direction vector of the plane spanned by $\gamma(t_1),\gamma(t_2),\gamma(t_3)$, it can be expressed as a linear combination of the form

$$\tilde{\mathbf{m}} = \lambda\ \gamma_1(t_1,t_2) + \mu\ \gamma_2(t_1,t_2,t_3).$$

The length of $\tilde{\mathbf{m}}$ tends to 1, $\gamma_1(t_1,t_2)$ and $\gamma_2(t_1,t_2,t_3)$ converge to two linearly independent vectors, thus, for any sequence of triples t_1,t_2,t_3 tending to $\bar{t}$, the corresponding sequence of λ's and μ's remain bounded. On the other hand,

$$\lim <\gamma_1(t_1,t_2),\mathbf{t}_i(\bar{t})> = <\gamma'(\bar{t}),\mathbf{t}_i(\bar{t})> = 0,$$
$$\lim <\gamma_2(t_1,t_2,t_3),\mathbf{t}_i(\bar{t})> = \frac{1}{2} <\gamma''(\bar{t}),\mathbf{t}_i(\bar{t})> = 0,$$

which implies

$$\lim \alpha_i = \lim <\mathbf{m},\mathbf{t}_i(\bar{t})> = \lim <\tilde{\mathbf{m}},\mathbf{t}_i(\bar{t})> = 0 \quad \text{for } i \geqslant 3.$$

c) We have $\alpha_2 = \langle \mathbf{m}, \mathbf{t}_2(\bar{t})\rangle \geqslant 0$ and $\| \mathbf{m} \|^2 = \sum \alpha_i^2 = 1$ by our assumptions. By a) and b), α_2 must tend to 1. This finishes the proof of the Lemma.

Substituting $\mathbf{p} = \gamma(\xi_1) + \rho\,\mathbf{m}$ into (**) we get

$$\langle \gamma''(\eta), \gamma(\eta) - \gamma(\xi_1) - \rho\,\mathbf{m} + \langle \gamma'(\eta), \gamma'(\eta)\rangle = 0,$$

from which

$$\rho = \frac{\langle \gamma''(\eta), \gamma(\eta) - \gamma(\xi_1)\rangle + \langle \gamma'(\eta), \gamma'(\eta)\rangle}{\langle \gamma''(\eta), \mathbf{m}\rangle}.$$

Since ξ_1, ξ_2 and η are sandwiched between t_1 and t_3, they converge to $\bar{t}$ if t_1, t_2 and t_3 tend to $\bar{t}$. In this case $\mathbf{p}$ tends to the vector

$$\gamma(\bar{t}) + \frac{\langle \gamma''(\bar{t}), \gamma(\bar{t}) - \gamma(\bar{t})\rangle + \langle \gamma'(\bar{t}), \gamma'(\bar{t})\rangle}{\langle \gamma''(\bar{t}), \mathbf{t}_2(\bar{t})\rangle}\,\mathbf{t}_2(\bar{t}) = \gamma(\bar{t}) + \frac{1}{\kappa_1(\bar{t})}\,\mathbf{t}_2(\bar{t}).$$

The radius of the circle $C(t_1, t_2, t_3)$ is the distance between $\gamma(t_1)$ and $\mathbf{p}$. This will obviously tend to the distance between $\gamma(\bar{t})$ and $\gamma(\bar{t}) + \frac{1}{\kappa_1(\bar{t})}\,\mathbf{t}_2(\bar{t})$, i.e. to $\frac{1}{|\kappa_1(\bar{t})|}$ as t_1, t_2, t_3 tend to $\bar{t}$. ▮

From now on we restrict our attention to plane curves.

Definition. The center of the osculating circle is called the center of curvature, the radius of the osculating circle is called the radius of curvature of the given curve at the given point.

Definition. The locus of the centers of curvature of a curve is called the evolute of the curve. The evolute is defined for arcs along which the curvature is not zero.

If $\gamma : [a,b] \longrightarrow \mathbb{R}^2$ is a curve with a nowhere zero curvature function κ and unit normal vector field $\mathbf{n}$, then the evolute can be parameterized by the mapping $\tilde{\gamma} : [a,b] \longrightarrow \mathbb{R}^2$, $\tilde{\gamma} = \gamma + (1/\kappa)\,\mathbf{n}$.

Exercise. Show that the evolute of the ellipse $\gamma(t) = (a \cos t, b \sin t)$ is the "affine astroid" $\tilde{\gamma}(t) = \left(\frac{a^2-b^2}{a} \cos^3 t\ ,\ \frac{b^2-a^2}{b} \sin^3 t \right)$.

The evolute of a curve was introduced by Huygens in connection with his investigations on the propagation of wave fronts. If we generate a curvilinear wave on the surface of calm water (e.g. we drop a wire into it), the wave starts moving. Mathematically, consecutive positions of the wave front are described by the parallel curves of the original curve.

Definition. Let γ be a regular plane curve with normal vector field $\mathbf{n}$. A parallel curve of γ is a curve of the form $\gamma_d = \gamma + d\,\mathbf{n}$, where $d \in \mathbb{R}$ is a fixed real.

Experiments on wave fronts show that even if the initial wave front is a smooth curve, singularities may appear on the wave front during its motion which move for a while and then disappear.

Definition. Let γ be a smooth parameterized curve, t a point of its domain. We say that $\gamma(t)$ (or t) is a singular point (or singular parameter) of the curve γ if $\gamma'(t) = \mathbf{0}$.

Proposition. Singular points of the parallel curves of a regular curve γ sweep out the evolute of γ.

Proof. Since $\gamma_d' = \gamma' + d\ \mathbf{n}' = \upsilon\ \mathbf{t} - \upsilon\ d\ \kappa\ \mathbf{t} = (1 - d\ \kappa)\ \upsilon\ \mathbf{t}$, singular parameters are characterized by $1 - d\ \kappa(t) = 0$. Then the corresponding singular points $\gamma_d(t) = \gamma(t) + (1/\kappa(t))\ \mathbf{n}(t)$ lie on the evolute of the curve. It is also easy to show that any evolute point is a singular point of a suitable parallel curve.█

Exercise. Study the singularities on the parallel curves of an ellipse.

Proposition. Let $\gamma : [a,b] \longrightarrow \mathbb{R}^2$ be a regular plane curve with curvature $\kappa \neq 0$ such that $\kappa' > 0$, and evolute $\tilde{\gamma} = \gamma + (1/\kappa)\ \mathbf{n}$. Then the normal of γ at $t \in [a,b]$ is tangent to the evolute at $\tilde{\gamma}(t)$ and the length of the arc of the evolute between $\tilde{\gamma}(t_1)$ and $\tilde{\gamma}(t_2)$, $t_1 < t_2$, is the difference of the radii of curvatures $1/\kappa(t_1) - 1/\kappa(t_2)$.

Proof. The speed vector of $\tilde{\gamma}$ is

$$\tilde{\gamma}' = \gamma' + (1/\kappa)'\ \mathbf{n} + (1/\kappa)\ \mathbf{n}' = \upsilon\ \mathbf{t} + (1/\kappa)'\ \mathbf{n} - (1/\kappa)\ \upsilon\ \kappa\ \mathbf{t} = (1/\kappa)'\ \mathbf{n}.$$

This equation shows that $\tilde{\gamma}$ is a regular curve and its tangent at $\tilde{\gamma}(t)$ is parallel to the normal $\mathbf{n}(t)$ of γ, which proves the first part of the proposition. As for the length of the evolute, it is equal to the integral

$$\int_{t_1}^{t_2} |\tilde{\gamma}(\tau)|\ d\tau = \int_{t_1}^{t_2} |(1/\kappa)'(\tau)\ \mathbf{n}(\tau)|\ d\tau = \int_{t_1}^{t_2} |(1/\kappa)'(\tau)|\ d\tau = \int_{t_1}^{t_2} -(1/\kappa)'(\tau)\ d\tau =$$

$= 1/\kappa(t_1) - 1/\kappa(t_2)$. █

The above proposition gives a method to construct the curve γ from its evolute. Suppose for simplicity that $\kappa > 0$. Take a thread of length $1/\kappa(a)$ and fix one of its ends to $\tilde{\gamma}(a)$. Then pulling the other end of the thread wrap it on the curve $\tilde{\gamma}$ starting from $\gamma(a)$. By the proposition, the moving end of the thread will slip along γ. Mathematically, the thread construction gives an involute of a curve.

Definition. Let $\gamma : [a,b] \longrightarrow \mathbb{R}^2$ be a unit speed curve with unit tangent vector field $\mathbf{t}$. An involute of the curve γ is a curve $\hat{\gamma}$ of the form $\hat{\gamma}(s) = \gamma(s) + (\ell - s)\ \mathbf{t}(s)$, where ℓ is a given real number.

A curve has many involutes corresponding to the different choices of the length ℓ of the thread.

Corollary. A curve satisfying the conditions of the previous proposition is an involute of its evolute (more exactly a reparameterization of it).

Proposition. Let γ be a unit speed curve with $\kappa > 0$, $\hat{\gamma}(s) = \gamma(s)+(\ell-s)\mathbf{t}(s)$ an involute of it such that ℓ is greater than the length of γ. Then the evolute of $\hat{\gamma}$ is γ.

Proof. We have

$$\hat{\gamma}'(s) = \mathbf{t}(s) - \mathbf{t}(s) + (\ell-s)\,\kappa(s)\,\mathbf{n}(s) = (\ell-s)\,\kappa(s)\,\mathbf{n}(s),$$
$$\hat{\gamma}''(s) = [(\ell-s)\,\kappa(s)]'\,\mathbf{n}(s) - (\ell-s)\,\kappa^2(s)\,\mathbf{t}(s).$$

The first equation implies that the Frenet frame $\hat{\mathbf{t}},\hat{\mathbf{n}}$ of $\hat{\gamma}$ is related to that of γ by $\hat{\mathbf{t}} = \mathbf{n}$, $\hat{\mathbf{n}} = -\mathbf{t}$. Computing the curvature $\hat{\kappa}$ of $\hat{\gamma}$,

$$\hat{\kappa}(s) = \frac{\langle \hat{\gamma}''(s), \hat{\mathbf{n}}(s)\rangle}{|\hat{\gamma}'(s)|^3} = \frac{(\ell-s)^2\kappa^3(s)}{(\ell-s)^3\kappa^3(s)} = \frac{1}{(\ell-s)} .$$

Thus, the evolute of $\hat{\gamma}$ is

$$\hat{\gamma} + (1/\hat{\kappa})\,\hat{\mathbf{n}} = \gamma + (\ell-s)\,\mathbf{t} - (\ell-s)\,\mathbf{t} = \gamma . \blacksquare$$

We formulate some further results on the evolute and involute as an exercise.

Exercise.

a) Suppose that the regular curves γ_1 and γ_2 have regular evolutes. Show that γ_1 and γ_2 are parallel if and only if their evolutes are the same.

b) Show that if two involutes of a regular curve are regular, then they are parallel.

Let $\gamma: [a,b] \longrightarrow \mathbb{R}^2$ be a *unit speed* curve. The direction angle $\alpha(s)$ of the tangent $\mathbf{t}(s)$ is determined only up to an integer multiple of 2π, however one can see easily the existence of a *differentiable* function $\alpha : [a,b] \longrightarrow \mathbb{R}$ such that $\alpha(s)$ is a direction angle of $\mathbf{t}(s)$ for all $s \in [a,b]$. Then

$$\mathbf{t}(s) = \big(\cos\alpha(s), \sin\alpha(s)\big).$$

Differentiating with respect to s,

$$\dot{\mathbf{t}} = \dot{\alpha}(-\sin\alpha, \cos\alpha).$$

If we compare this equality with Frenet equations, we see immediately that

$$\kappa = \dot{\alpha} ,$$

i.e. the curvature is the derivative of the direction angle of the tangent vector with respect to the arc length.

Definition. The total curvature of a curve is the integral of its curvature function with respect to arc length

$$\int_a^b \kappa(s)\,ds .$$

By the relation $\kappa = \dot{\alpha}$, the total curvature of a curve is $\alpha(b) - \alpha(a)$, thus it measures the rotation made by the tangent vector during the motion along the curve from the initial point to the end of it.

Definition. A curve $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ is a smooth closed curve, if there exists a smooth mapping $\tilde{\gamma} : \mathbb{R} \longrightarrow \mathbb{R}^n$ such that $\tilde{\gamma}|_{[a,b]} = \gamma$, and $\tilde{\gamma}$ is periodic with period b-a $\tilde{\gamma}(t + b - a) \equiv \tilde{\gamma}(t)$.

In other words, a curve is closed if it returns to its initial point in such a way that arriving at the end of the curve one can smoothly go through from the end to the beginning and thus continue the motion periodically until infinity. Since the tangent vector of a smooth closed curve is the same at the endpoints $\gamma(a)$ and $\gamma(b)$, the direction angles $\alpha(b)$ and $\alpha(a)$ differ in an integer multiple of 2π.

Definition. The integer $\big(\alpha(b)-\alpha(a)\big)/2\pi$ is called the rotation number of the closed curve γ.

Exercise. Construct a closed curve with an arbitrarily given rotation number $k \in \mathbb{Z}$.

Solving the exercise we may see that all our efforts to construct a curve with rotation number $\neq \pm 1$ having no self-intersection are in vain. The reason for this is the famous "Umlaufsatz" (rotation number theorem).

Definition. A curve $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ is called simple if it has no self-intersection, i.e. $\gamma(t) \neq \gamma(t')$ whenever $t \neq t'$; γ is a simple closed curve if it is closed and we may have $\gamma(t) = \gamma(t')$ for $t \neq t'$ only in the case $\{t,t'\} = \{a,b\}$.

Theorem. (Umlaufsatz) The rotation number of a simple closed curve in the plane is equal to ± 1, or equivalently the total curvature of a simple closed curve is $\pm 2\pi$ hence independent of the actual shape of the curve!

The proof of theorem will use some new concepts borrowed from topology.

Definition. Let S^1 denote the unit circle $\{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x}| = 1\}$. The mapping $\pi : \mathbb{R} \longrightarrow S^1$ $\pi(t) = (\cos t, \sin t)$ is called the universal covering map of S^1. Given a continuous mapping $\varphi : X \longrightarrow S^1$ from a topological space to S^1, we say that φ has a lifting to $\mathbb{R}$ if there exist a continuous mapping $\bar{\varphi} : X \longrightarrow \mathbb{R}$ such that $\varphi = \bar{\varphi} \circ \pi$.

Exercise. Construct a continuous mapping from a topological space into the circle which has no lifting.

Lemma. Suppose that the image of the mapping $\phi : X \longrightarrow S^1$ does not cover the point $(\cos \alpha, \sin \alpha) \in S^1$ and that the restriction φ of ϕ onto a subspace

$Y \subset X$ has got a lifting $\bar{\varphi}$ such that $\bar{\varphi}(Y)$ is contained in an interval of the form $(\alpha+2k\pi, \alpha+2(k+1)\pi)$, $k \in \mathbb{Z}$. Then $\bar{\varphi}$ can be extended to a lifting $\bar{\phi}$ of ϕ. If furthermore Y is non-empty, X is path connected (i.e. any two points of X can be connected by a continuous curve lying in X) then the lifting $\bar{\phi}$ is unique, and maps X into the interval $(\alpha+2k\pi, \alpha+2(k+1)\pi)$.

Proof. The restriction of π onto $(\alpha+2k\pi, \alpha+2(k+1)\pi)$ is a homeomorphism between $(\alpha+2k\pi, \alpha+2(k+1)\pi)$ and $S^1 \setminus \{(\cos\alpha, \sin\alpha)\}$. Thus $\bar{\phi}$ can be defined as $\bar{\phi} = (\pi|_{(\alpha+2k\pi, \alpha+2(k+1)\pi)})^{-1} \circ \phi$. If X is path connected, then so is its image under a continuous lifting $\bar{\phi}$. Consequently $\bar{\phi}(X)$ must be contained in one of the intervals $(\alpha+2k\pi, \alpha+2(k+1)\pi)$. If Y is non-empty then k is uniquely determined, hence $\bar{\phi}$ must have the form $(\pi|_{(\alpha+2k\pi, \alpha+2(k+1)\pi)})^{-1} \circ \phi$. ▌

Proposition. Any continuous mapping $\varphi : [a,b] \longrightarrow S^1$ from an interval into the circle has got a lifting.

Proof. Choose a partition $a = t_0 < t_1 < \ldots < t_k = b$ of the interval $[a,b]$ fine enough to insure that the restriction of φ onto $[t_i, t_{i+1}]$ does not cover the whole circle. This is possible because of the continuity of φ. Then using the lemma we can define a lifting $\bar{\varphi}$ of φ step by step, extending $\bar{\varphi}$ recursively to $[a, t_i]$, $i=1,2,\ldots,k$. ▌

Remark. If φ is differentiable, then so is its lifting.

Proposition. Any continuous mapping $\varphi : T \longrightarrow S^1$ from a rectangle $T = [a,b] \times [c,d]$ into the circle has got a lifting to $\mathbb{R}$.

Proof. Divide the rectangle into nxn small rectangles so that the image of any of the small rectangles does not cover the circle. Then we may define a lifting $\bar{\varphi}$ of φ recursively, applying at each step the lemma above. $\bar{\varphi}$ can be defined first on the small rectangles of the first row going from left to right then on the rectangles of the second row, etc.

1	2	3	...	n
n+1	n+2	n+3	...	2n
	.	. .	.	

Proof of the "Umlaufsatz". Let us choose a point p on the regular simple closed curve γ in such a way that the curve is contained on one side of the tangent at p and parameterize the curve by arc length starting from p. Denoting this parameterization also by $\gamma : [0,\ell] \longrightarrow \mathbb{R}^2$, we define a mapping $\varphi : [0,\ell] \times [0,\ell] \longrightarrow S^1$ by

$$\varphi(t_1,t_2) = \begin{cases} -\gamma'(0)/\|\gamma'(0)\| & \text{if } t_1=\ell,\ t_2=0, \\ \gamma'(0)/\|\gamma'(0)\| & \text{if } t_1=0,\ t_2=\ell, \\ \dfrac{\gamma(t_1)-\gamma(t_2)}{\|\gamma(t_1)-\gamma(t_2)\|} & \text{if } t_1>t_2 \text{ and } \{t_1,t_2\}\neq\{0,\ell\}, \\ \gamma'(t_1)/\|\gamma'(t_1)\| & \text{if } t_1=t_2, \\ \dfrac{\gamma(t_2)-\gamma(t_1)}{\|\gamma(t_2)-\gamma(t_1)\|} & \text{if } t_2>t_1 \text{ and } \{t_1,t_2\}\neq\{0,\ell\}. \end{cases}$$

It is easy to see that φ is continuous, so it has a continuous lifting $\bar{\varphi}$. If the function $\alpha: [0,\ell] \longrightarrow \mathbb{R}$ is defined by $\alpha(t) = \bar{\varphi}(t,t)$, then $\alpha(t)$ is a direction angle of the speed vector $\mathbf{t}(t)$ of γ thus $\alpha(\ell)-\alpha(0)$ is 2π times the rotation number of γ. Consider the functions $\xi(t) = \bar{\varphi}(0,t)$ and $\vartheta(t) = \bar{\varphi}(t,\ell)$. $\xi(t)$ is a direction angle of the unit vector $\gamma(t)-p$ / $\|\gamma(t)-p\|$, $\vartheta(t)$ is a direction angle of its opposite. Thus ξ and ϑ differ only in a constant of the form $(2k+1)\pi$. Since the vectors $\gamma(t)-p$ / $\|\gamma(t)-p\|$ point in a half-plane bounded by the tangent at p, the image of ξ is contained in an open interval of length 2π (see lemma). Thus $\xi(\ell)-\xi(0)$, which has obviously the form $(2m+1)\pi$ for some $m \in \mathbb{Z}$, must be equal to $\pm\pi$. Hence, we conclude that

$$\alpha(\ell)-\alpha(0) = \bar{\varphi}(\ell,\ell)-\bar{\varphi}(0,0) = \big(\bar{\varphi}(\ell,\ell)-\bar{\varphi}(0,\ell)\big) + \big(\bar{\varphi}(0,\ell)-\bar{\varphi}(0,0)\big) =$$
$$= \vartheta(\ell)-\vartheta(0) + \xi(\ell) - \xi(0) = 2\big(\xi(\ell) - \xi(0)\big) = \pm 2\pi. \quad ▌$$

Remark. With more work but using essentially the same idea, one can generalize the "Umlaufsatz" for piecewise smooth closed simple curves. The generalization says that for a simple closed polygon with smooth curvilinear edges, the sum of oriented external angles plus the sum of the total curvatures of the edges equals $\pm 2\pi$.

Definition. A simple closed curve γ is convex, if for any point $P = \gamma(\bar{t})$, the curve lies on one side of the tangent to γ at P. In other words the function $\langle\gamma(t),\underline{n}(\bar{t})\rangle$ must be $\geqslant 0$ or $\leqslant 0$ for all t.

Exercise. Show that a simple closed curve is convex if and only if every arc of the curve lies on one side of the straight line through the endpoints of the arc.

Convex curves can be characterized with the help of the curvature function.

Theorem. A simple closed curve is convex if and only if $\kappa \geqslant 0$ or $\kappa \leqslant 0$ everywhere along the curve.

Proof. Assume first that γ is a naturally parameterized convex curve. Let $\alpha(t)$ be a continuous direction angle for the tangent $\underline{t}(t)$. As we know,

$\alpha^{\cdot} = \kappa$, thus, it suffices to show that α is a weakly monotonous function. This follows if we show that if α takes the same value at two different parameters t_1, t_2, then α is constant on the interval $[t_1, t_2]$.

The rotation number of a simple curve is ±1, hence the image of $\underline{t}$ covers the whole unit circle. As a consequence, we can find a point at which

$$\underline{t}(t_3) = -\underline{t}(t_1) = -\underline{t}(t_2).$$

If the tangent lines at t_1, t_2, t_3 were different, then one of them would be between the others and this tangent would have points of the curve on both sides. This contradicts convexity, hence two of these tangents say the tangents at $P = \gamma(t_i)$ and $Q = \gamma(t_j)$ coincide.

We claim that the segment $\overline{PQ}$ is an arc of γ. It is enough to prove that this segment is in the image of γ. Assume to the contrary that a point $X \in \overline{PQ}$ is not covered by γ. Drawing a line $e \neq PQ$ through X, we can find at least two intersection points R and S of e and the curve, since e separates P and Q and γ has two essentially disjoint arcs connecting P to Q. Since PQ is a tangent of γ, the point R and S must lie on the same side of it. As a consequence, we get that one of the triangles PQR and PQS, say the first one is inside the other. However, this leads to a contradiction, since for such a configuration the tangent through S necessarily separates two vertices of the triangle PQR, which lie lie on the curve.

If γ is defined on the interval $[a,b]$, then $\gamma(a) = \gamma(b)$ is either on the segment $\overline{PQ}$ or not. The first case is not possible, because then α would be constant on the intervals $[a, t_1]$ and $[t_2, b]$, yielding

$$\alpha(a) = \alpha(t_1) = \alpha(t_2) = \alpha(b)$$

and

$$\text{rotation number} = (\alpha(b) - \alpha(b))/2\pi = 0.$$

In the second case α is constant on the interval $[t_1, t_2]$, as we wanted to show.

Now to prove the converse, assume that γ is a simple closed curve with $\kappa \geqslant 0$ everywhere and assume to the contrary that γ is not convex (the case $\kappa \leqslant 0$ can be treated analogously). Then we can find a point $P = \gamma(t_1)$, such that the tangent at P has curve points on both of its sides. Let us find on each side a curve point, say $Q = \gamma(t_2)$ and $R = \gamma(t_3)$ respectively, lying at maximal distance from the tangent at P. Then the tangents at P, Q and R are different and parallel. Since the unit tangent vectors $\underline{t}(t_i)$, $i=1,2,3$ have parallel directions, two of them, say $\underline{t}(t_i)$ and $\underline{t}(t_j)$ must be equal. $A = \gamma(t_i)$ and $B = \gamma(t_j)$ divide the curve into two arcs. Denoting by K_1 and K_2

the total curvatures of these arcs, we deduce that these total curvatures have the form $K_1 = 2k_1\pi$, $K_2 = 2k_2\pi$, where $k_1, k_2 \in \mathbb{Z}$, since the unit tangents at the ends of the arcs are equal. On the other hand, we have $k_1+k_2 = 1$ by the Umlaufsatz and $k_1 \geqslant 0$, $k_2 \geqslant 0$ by the assumption $\kappa \geqslant 0$. This is possible only if one of the total curvatures K_1 or K_2 is equal to zero. Since $\kappa \geqslant 0$, this means that $\kappa = 0$ along one of the arcs between A and B. But then this arc would be a straight line segment, implying that the tangents at A and B coincide. The contradiction proves the theorem. ▮

Definition. A point $\gamma(t)$ of a regular plane curve γ is called a vertex if $\kappa'(t) = 0$.

Vertices of a curve correspond to the singular points of the evolute.

By compactness, the curvature function of a closed curve attains somewhere its maximum and minimum, hence every closed curve has at least two vertices.

Exercise. Find a parameterization of Bernulli's lemniscate

$$\{P \in \mathbb{R}^2 : \overline{PA}\,\overline{PB} = 1/4\,(\overline{AB})^2\},$$

where $A \neq B$ are given points in the plane, plot the curve and show that it is a closed curve with exactly two vertices. Determine the rotation number of the lemniscate.

Theorem (Four Vertex Theorem). A convex closed curve has at least 4 vertices.

This result is sharp, since an ellipse has exactly four vertices.

Proof. Local maxima and minima of the curvature function yield vertices. One can always find a local minimum on an arc bounded by two local maxima, hence if we have two local maxima or minima of the curvature then we must have at least four vertices. Thus we have to exclude the case when the curvature function has one absolute maximum at A and one absolute minimum at B, and strictly monotonous on the arcs bounded by A and B. In this case, choose a coordinate system with origin at A and x-axis AB.

The arcs of the curve bounded by A and B do not cut the x-axis at points other then A and B. Indeed, if there were a further intersection point C, then the curve would be split into three arcs by A,B and C in such a way that on each arc we could find a point at which the tangent to the curve is parallel to the straight line ABC. If the three tangent at these points were different then the one in the middle would cut the curve apart contradicting to convexity, if two of the tangents coincided then we could find a straight line segment contained in the curve yielding an infinite number of vertices.

If the two arcs bounded by A and B lied on the same side of AB then the line AB would be a common tangent of the curve at A and B. In this case the segment $\overline{AB}$ would be contained in the curve, yielding an infinite number of vertices as before.

We conclude that for a suitable orientation of the y-axis, $y(t)\kappa'(t) \geq 0$ for every $t \in [a,b]$, where $\gamma(t) = (x(t),y(t))$ $t \in [a,b]$ is a unit speed parameterization of the curve. Hence we get

$$\int_a^b y(t)\kappa'(t)dt > 0.$$

Integrating by parts,

$$\int_a^b y(t)\kappa'(t)dt = \Big[y(t)\kappa(t)\Big]_a^b - \int_a^b y'(t)\kappa(t)dt = -\int_a^b y'(t)\kappa(t)dt.$$

The unit tangent vector field of the curve is $\underline{t} = (x',y')$, the unit normal vector field is $\underline{n} = (-y',x')$, hence by the first Frenet formula,

$$x'' = -\kappa y'.$$

Integrating,

$$-\int_a^b y'(t)\kappa(t)dt = \int_a^b x''(t)dt = \Big[x''(t)\Big]_a^b = 0.$$

This is a contradiction since a positive number can not be equal to 0. ∎

Further Exercises

3-1. Find the points on the ellipse $\gamma(t) = (a \cos t, b \sin t)$ at which the curvature is minimal or maximal $(a > b > 0)$.

3-2. The curve "cardioid" is the trajectory of a peripheral point of a circle rolling about a fixed circle of the same radius.

- Find a smooth parameterization of the cardioid.
- Compute its length.
- Show that its evolute is also a cardioid.

3-3. The "chain curve" is the graph of the hyperbolic cosine function

$$ch(x) = \frac{e^x + e^{-x}}{2}.$$

- Determine the involute of the chain curve touching the chain curve at (0,1). (This curve is called "tractrix".)

- Let the tangent of the tractrix at P intersect the x-axis at Q. Show that the segment PQ has unit length.

3-4. Let γ be a simple regular closed curve of length ℓ with curvature function κ. Choose a real number d such that $1 \geqslant \kappa\, d$. How long is the parallel curve $\gamma_d = \gamma + d\,\mathbf{n}$?

3-5. Let γ be a regular plane curve for which the curvature function and its derivative are positive. Show that for any $t_1 < t_2$ from the parameter domain of γ the osculating circle of γ at $\gamma(t_1)$ contains the osculating circle at $\gamma(t_2)$.

UNIT 4. 3D CURVES - CURVES ON HYPERSURFACES

Explicit formulas, projections of a space curve onto the coordinate planes of the Frenet basis, the shape of a curve around one of its points, hypersurfaces, regular hypersurface, tangent space and unit normal of a hypersurface, curves on hypersurfaces, normal sections, normal curvatures, Meusnier's theorem.

A 3-dimensional curve is a curve of general type if its first two derivatives are not parallel. From now on we shall suppose that the curves in question are all of general type.

The distinguished Frenet frame vector fields $\mathbf{t}_1, \mathbf{t}_2$ and $\mathbf{t}_3$ of a 3-dimensional curve are denoted in classical differential geometry by $\mathbf{t}$, $\mathbf{n}$ and $\mathbf{b}$ and they are called the *(unit) tangent*, the *principal normal* and the *binormal* vector fields of the curve respectively. These vector fields define a coordinate system at each point of the curve. The coordinate planes of this coordinate system are given the following names. We are already familiar with the plane that goes through a given curve point and spanned by the directions of the tangent and principal normal. It is the *osculating plane* of the curve. The plane that is spanned by the principal normal and the binormal is the plane that contains all straight lines orthogonally intersecting the curve at the given point. For this obvious reason, this plane is called the *normal plane* of the curve. The third coordinate plane, that is the plane spanned by the tangent and the binormal directions is the *rectifying plane* of the curve. The reason for this naming will become clear later. As we know from the general theory, a 3D curve of general type has two curvature functions κ_1, which is always positive and κ_2, which may have any sign. The first curvature κ_1 is denoted by the classics simply by κ and it is referred to as the *curvature* of the curve while the second curvature κ_2 is called the *torsion* of the curve and is denoted by τ.

Using the classical notation, Frenet formulas for a space curve can be written as follows.

$$\mathbf{t}' = v\,\kappa\,\mathbf{n}$$
$$\mathbf{n}' = -v\,\kappa\,\mathbf{t} + v\,\tau\,\mathbf{b}$$
$$\mathbf{b}' = -v\,\tau\,\mathbf{n}.$$

Now let us find explicit formulas for the computation of these vectors and curvatures in an economic way. The formulas we shall derive involve the "cross product" of vectors. Let us recall the definition and basic properties of this operation. The cross product of two vectors can be defined in a geometric and in an algebraic way. According to the geometric definition, the cross product $\mathbf{a}\times\mathbf{b}$ of the vectors $\mathbf{a}$ and $\mathbf{b}$ is $\mathbf{0}$ if $\mathbf{a}$ and $\mathbf{b}$ are parallel; if $\mathbf{a}$ and $\mathbf{b}$ are not parallel, then it is the vector defined by the following three conditions

i) $\mathbf{a}\times\mathbf{b}$ is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$;

ii) $\|\mathbf{a}\times\mathbf{b}\|$ is equal to the area of the parallelogram spanned by $\mathbf{a}$ and $\mathbf{b}$;

iii) $(\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b})$ is a positively oriented (right handed) basis of $\mathbb{R}^3$.

Algebraically we can introduce the cross product in the following way. Let $\mathbf{a} = (a_1,a_2,a_3)$, $\mathbf{b} = (b_1,b_2,b_3)$; $\mathbf{e}_1 = (1,0,0)$ $\mathbf{e}_2 = (0,1,0)$, $\mathbf{e}_3 = (0,0,1)$. The cross product of $\mathbf{a}$ and $\mathbf{b}$ is the determinant

$$\mathbf{a}\times\mathbf{b} = \det \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (\ a_2b_3-a_3b_2\ ,\ a_3b_1-a_1b_3\ ,\ a_1b_2-a_2b_1\ ).$$

Here are some basic properties of the cross product.

i) $\mathbf{a}\times\mathbf{b} = \mathbf{0}$ if and only if $\mathbf{a}$ and $\mathbf{b}$ are parallel;

ii) $\mathbf{a}\times\mathbf{b} = -\ \mathbf{b}\times\mathbf{a}$;

iii) $(\mathbf{a}+\mathbf{b})\times\mathbf{c} = \mathbf{a}\times\mathbf{c} + \mathbf{b}\times\mathbf{c}$;

iv) $\mathbf{c}\times(\mathbf{a}+\mathbf{b}) = \mathbf{c}\times\mathbf{a} + \mathbf{c}\times\mathbf{b}$;

v) $(\lambda\mathbf{a})\times\mathbf{b} = \mathbf{a}\times(\lambda\mathbf{b}) = \lambda(\mathbf{a}\times\mathbf{b}) \quad \lambda\in\mathbb{R}$.

<u>Exercise</u>. Show the equivalence of the geometric and algebraic definitions and prove basic properties of cross product.

Now let $\gamma:[a,b]\longrightarrow \mathbb{R}^3$ be a curve of general type. The unit tangent vector field $\mathbf{t}$ can be obtained by normalizing the speed vector γ'

$$\mathbf{t} = \frac{\gamma'}{\|\gamma'\|} \quad .$$

To obtain the principal normal $\mathbf{n}$ we can use the general method based on Gram Schmidt orthogonalization process

$$\mathbf{n} = \frac{\gamma'' - \langle\gamma'',\mathbf{t}\rangle\ \mathbf{t}}{\|\gamma'' - \langle\gamma'',\mathbf{t}\rangle\ \mathbf{t}\|} = \frac{\|\gamma'\|^2\ \gamma'' - \langle\gamma'',\gamma'\rangle\ \gamma'}{\|\ \|\gamma'\|^2\ \gamma'' - \langle\gamma'',\gamma'\rangle\ \gamma'\ \|} \quad ,$$

and after this we can compute the binormal as a cross product of $\mathbf{t}$ and $\mathbf{n}$

$$\mathbf{b} = \mathbf{t} \times \mathbf{n} \ .$$

In practice however, it is more convenient to calculate the binormal first. The binormal vector is the unit normal vector of the osculating plane for which $(\mathbf{t},\mathbf{n},\mathbf{b})$ is positively oriented. The osculating plane is spanned by the first two derivatives γ', γ'' of γ, furthermore the pair $(\mathbf{t}, \mathbf{n})$ defines the same orientation of the osculating plane as the pair (γ', γ'') so the basis $(\gamma', \gamma'', \mathbf{b})$ is positively oriented. Hence,

$$\mathbf{b} = \frac{\gamma' \times \gamma''}{\|\gamma' \times \gamma''\|} .$$

Having computed $\mathbf{b}$, $\mathbf{n}$ can be obtained as

$$\mathbf{n} = \mathbf{b} \times \mathbf{t} = \frac{(\gamma' \times \gamma'') \times \gamma'}{\|\gamma' \times \gamma''\| \; \|\gamma'\|} .$$

Before the computation of the curvature and torsion let us express the first three derivatives of γ as linear combinations of the Frenet vectors.

$$\gamma' = v\, \mathbf{t}$$

$$\gamma'' = v'\, \mathbf{t} + v\, \mathbf{t}' = v'\, \mathbf{t} + v^2 \kappa\, \mathbf{n}$$

$$\begin{aligned}\gamma''' &= v''\, \mathbf{t} + v' \mathbf{t}' + (v^2 \kappa)' \mathbf{n} + v^2 \kappa\, \mathbf{n}' = \\ &= v''\, \mathbf{t} + v' v \kappa\, \mathbf{n} + (v^2 \kappa)' \mathbf{n} + v^2 \kappa\, v\, (-\kappa\, \mathbf{t} + \tau\, \mathbf{b}) = \\ &= \left(v'' - v^3 \kappa^2\right) \mathbf{t} + \left(v' v \kappa + (v^2 \kappa)'\right) \mathbf{n} + \left(v^3 \kappa \tau\right) \mathbf{b} .\end{aligned}$$

From the first two equations

$$\gamma' \times \gamma'' = v\, \mathbf{t} \times \left(v'\, \mathbf{t} + v^2 \kappa\, \mathbf{n}\right) = v^3 \kappa\, \mathbf{b} .$$

Taking the length of these vectors and using that v and κ are positive,

$$\| \gamma' \times \gamma'' \| = v^3 \kappa ,$$

from which

$$\boxed{\kappa = \frac{\| \gamma' \times \gamma'' \|}{\| \gamma' \|^3}}$$

The torsion of the curve is involved only in the coefficient of $\mathbf{b}$ in the expression for γ'''. We can draw out the essential information for the torsion and get rid of the "rubbish" by taking the dot product of this expression with $\mathbf{b}$ or a vector parallel with $\mathbf{b}$. Since $\gamma' \times \gamma'' \parallel \mathbf{b}$, we get

$$\langle \gamma' \times \gamma'', \gamma''' \rangle = v^6 \kappa^2 \tau .$$

Combining this equation with the expression we have for the length of $\gamma' \times \gamma''$,

$$\boxed{\tau = \frac{\langle \gamma' \times \gamma'', \gamma''' \rangle}{\| \gamma' \times \gamma'' \|^{2}}}$$

Recall that the numerator of this fraction is the determinant of the matrix the rows of which are $\gamma', \gamma'', \gamma'''$ and geometrically, it is the signed volume of the parallelepiped spanned by $\gamma', \gamma'', \gamma'''$, where the sign is positive if and only if $(\gamma', \gamma'', \gamma''')$ is positively oriented.

Now we are going to study the shape of the orthogonal projections of a curve onto the planes spanned by the vectors of the distinguished Frenet frame. For simplicity, suppose that the curve γ is parameterized by arc length and examine the curve around $\gamma(0)$. Since $\kappa \equiv 1$, the formulas that express the derivatives of γ in terms of Frenet vectors reduce to the form

$$\gamma' = \mathbf{t}$$

$$\gamma'' = \mathbf{t}' = \kappa\,\mathbf{n}$$

$$\gamma''' = \kappa'\mathbf{n} + \kappa\,\mathbf{n}' = \kappa'\mathbf{n} + \kappa\,(-\kappa\,\mathbf{t} + \tau\,\mathbf{b}) = -\,\kappa^2\,\mathbf{t} + \kappa'\,\mathbf{n} + \kappa\,\tau\,\mathbf{b}\ .$$

We can approximate the curve γ around $\gamma(0)$ by its Taylor expansion.

$$\gamma(t) = \gamma(0) + \gamma'(0)t + \frac{\gamma''(0)}{2}\,t^2 + \frac{\gamma'''(0)}{6}\,t^3 + o(t^3).$$

Recall that the "little oh" notation $o(t^3)$ is used in the following sense. If f,g, and h are functions defined around a given point a, then we write

$$f(t) = g(t) + o(h(t))$$

if $\frac{f(t) - g(t)}{h(t)}$ tends to zero as t tends to a. Though a is not involved in the equality it is generally clear from the context what it is. For example, in our case a = 0.

Expressing the derivatives of γ with the help of Frenet vectors we get

$$\gamma(t) - \gamma(0) =$$

$$= \Big(t - \kappa^2(0)\frac{t^3}{6}\Big)\ \mathbf{t}(0) + \Big(\kappa(0)\frac{t^2}{2} + \kappa'(0)\frac{t^3}{6}\Big)\ \mathbf{n}(0) + \Big(\kappa(0)\tau(0)\frac{t^3}{6}\Big)\ \mathbf{b}(0) + o(t^3)\ .$$

Looking at this expansion we conclude that the projection of the curve on the osculating plane is well approximated by the parabola $t\mathbf{t}(0) + \kappa(0)\frac{t^2}{2}\mathbf{n}(0)$, (observe that the curvature of this parabola at t = 0 is $\kappa(0)$), the projection onto the normal plane has locally the same shape as the semicubical parabola $\kappa(0)\frac{t^2}{2}\mathbf{n}(0) \ +\frac{t^3}{6}\big(\kappa'(0)\ \ \mathbf{n}(0)\ \ +\ \ \kappa(0)\tau(0)\ \ \mathbf{b}(0)\big)$, in particular, it has a so called cusp singularity at t = 0, finally, the projection onto the rectifying plane has the Taylor expansion

$\mathbf{t}(0) + \frac{t^3}{6}\ \big(\kappa^2(0)\kappa(0)\mathbf{t}(0)\ \tau(0)\ \mathbf{b}(0)\big) + o(t^3)$, so its shape is like the graph of a cubic function. It is easy to see, that the curvature of this projection is 0 at t = 0, thus it is almost straight around the origin. That is the reason why the rectifying plane was given just this name: projection of the curve onto the rectifying plane straightens the curve and "rectifying" means straightening.

Now we shall study a problem which connects curve theory to surface theory. If a curve lyes on a surface, curvedness of the surface forces the curve to bend. Thus, curvedness of a surface can be detected by the curvatures of the curves lying on the surface. Heuristically clear that the curvature of a curve on a given surface should be the same as the curvature of the intersection curve of the surface and the osculating plane of the curve provided that the osculating plane is not tangent to the surface. This is indeed true and thus we may pose the question how to compute the curvature of the curve using only information on the surface and the position of the osculating plane of the curve. The existence of a formula that answers this question will prove our heuristics.

Definition. A parameterized hypersurface in $\mathbb{R}^n$ is a differentiable mapping $\mathbf{r}:\Omega \longrightarrow \mathbb{R}^n$ from an open domain Ω of $\mathbb{R}^{n-1}$ into the n-dimensional space. Smooth curves on a parameterized hypersurface are curves of the form $\gamma(t) = \mathbf{r}(u(t))$, where the mapping $t \longmapsto u(t)$ is a smooth curve lying in the parameter domain Ω. Curves of the form $t \longmapsto \mathbf{r}(u_1,\dots,u_{i-1},t,u_{i+1},\dots,u_{n-1})$, where $u_1,\dots,u_{i-1},u_{i+1},\dots,u_{n-1}$ are fixed numbers are called the parameter lines or coordinate lines on the hypersurface. The speed vectors of the parameter lines $t \longmapsto \mathbf{r}(u_1,\dots,u_{i-1},t,u_{i+1},\dots,u_{n-1})$, which are just the partial derivatives of the mapping $\mathbf{r}$ with respect to the i-th variable, will be denoted by $\mathbf{r}_i(u_1,\dots,u_{i-1},t,u_{i+1},\dots,u_{n-1})$.

Since we shall often work with formulas containing partial derivatives of a function it is convenient to introduce the shorthand convention that we shall denote the partial derivative of a multivariable function F with respect to its i-th variable by F_i. In general, the higher order partial derivative $\frac{\partial^k F}{\partial u_{i_1}\dots\partial u_{i_k}}$ of F will be denoted by $F_{i_1\dots i_k}$. If there is a danger of confusion with lower indices, the lower indices of the function will be separated from the indices of variables with respect to which we have to take the partial derivative by a comma. Thus, $\frac{\partial^k F_{j_1\dots j_\ell}}{\partial u_{i_1}\dots\partial u_{i_k}}$ will be denoted by $F_{j_1\dots j_\ell,i_1\dots i_k}$.

Definition. A parameterized hypersurface is regular if the vectors $\mathbf{r}_1(u),\dots,\mathbf{r}_{n-1}(u)$ are linearly independent for any $u \in \Omega$. In this case we also say that $\mathbf{r}$ is an immersion of the domain Ω into $\mathbb{R}^n$.

Definition. The tangent plane of a regular parameterized hypersurface at the point $\mathbf{r}(u)$ is the plane through $\mathbf{r}(u)$ spanned by the direction vectors $\mathbf{r}_1(u),\dots,\mathbf{r}_{n-1}(u)$. The unit normal vector of the hypersurface at the point $\mathbf{r}(u)$ is defined to be the unit normal vector $\mathbf{N}(u)$ of the tangent plane, for which $\mathbf{r}_1(u),\dots,\mathbf{r}_{n-1}(u)$, $N(u)$ is a positively oriented basis of $\mathbb{R}^n$.

For parameterized surfaces in $\mathbb{R}^3$, the unit normal vector field can be calculated with the help of cross product

$$N(u_1,u_2) = \frac{\mathbf{r}_1(u_1,u_2) \times \mathbf{r}_2(u_1,u_2)}{\| \mathbf{r}_1(u_1,u_2) \times \mathbf{r}_2(u_1,u_2)\|}.$$

To get a similar formula in higher dimensions, we need a suitable generalization of the cross product.

Let $\mathbf{r}_i = (r_i^1,\dots,r_i^n) \in \mathbb{R}^n$, $i = 1,2,\dots,n-1$, be n-dimensional vectors, $\mathbf{e}_1,\dots,\mathbf{e}_n$ the standard basis of $\mathbb{R}^n$. The <u>exterior product</u> of the vectors $\mathbf{r}_1,\dots,\mathbf{r}_{n-1}$ is defined by the equality

$$\mathbf{r}_1\wedge\dots\wedge\mathbf{r}_{n-1} = \det \begin{vmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \\ r_1^1 & \dots & r_1^n \\ \dots & \dots & \dots \\ r_{n-1}^1 & \dots & r_{n-1}^n \end{vmatrix}.$$

<u>Exercise</u>. (cf. Ex. on page 18) Show that $\mathbf{r}_1\wedge\dots\wedge\mathbf{r}_{n-1}$ is orthogonal to $\mathbf{r}_1,\dots,\mathbf{r}_{n-1}$, it is different from $\mathbf{0}$ if and only if $\mathbf{r}_1,\dots,\mathbf{r}_{n-1}$ are linearly independent, and finally, $\mathbf{r}_1,\dots,\mathbf{r}_{n-1},(-1)^{n-1}\mathbf{r}_1\wedge\dots\wedge\mathbf{r}_{n-1}$ is a positively oriented basis of $\mathbb{R}^n$.

As a consequence of the exercise, for regular hypersurfaces we have

$$N(u) = (-1)^{n-1} \frac{\mathbf{r}_1(u)\wedge\dots\wedge\mathbf{r}_{n-1}(u)}{\| \mathbf{r}_1(u)\wedge\dots\wedge\mathbf{r}_{n-1}(u) \|}.$$

Consider the curve $\gamma(t) = \mathbf{r}(u(t))$ lying on the regular parameterized hypersurface $\mathbf{r}:\Omega\longrightarrow \mathbb{R}^n$, where $u = (u_1,\dots,u_{n-1})$ is a curve in Ω. Express the first two derivatives of γ using Frenet formulas on one hand and the special form of γ as a surface curve on the other. Using the chain rule, we get

$$\vartheta\, \mathbf{t}_1 = \gamma' = \sum_{i=1}^{n-1} u_i' \,\mathbf{r}_i(u)$$

and

$$\vartheta' \mathbf{t}_1 + \vartheta^2\kappa_1 \,\mathbf{t}_2 = \gamma'' = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} u_i'u_j' \,\mathbf{r}_{ij}(u) + \sum_{i=1}^{n-1} u_i'' \,\mathbf{r}_i(u).$$

Multiplying the last equation by the normal vector of the hypersurface and using the fact that it is orthogonal to the tangent vectors $\mathbf{t}_1,\mathbf{r}_1,\dots,\mathbf{r}_{n-1}$, we obtain

$$\vartheta^2\kappa_1 \langle\mathbf{N}(u),\mathbf{t}_2\rangle = \langle N(u),\gamma''\rangle = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \langle N(u),\mathbf{r}_{ij}(u)\rangle\, u_i'u_j',$$

from which

$$\boxed{\kappa_1 = \frac{1}{\langle\mathbf{N}(u),\mathbf{t}_2\rangle} \frac{\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \langle N(u),\mathbf{r}_{ij}(u)\rangle\, u_i'u_j'}{\vartheta^2}}$$

Let us study this expression. We claim that the right hand side is determined by the osculating plane of the curve and the surface provided that the osculating plane is not tangent to the surface.

Let us start with the expression $k(\gamma') = \dfrac{\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \langle N(u),\mathbf{r}_{ij}(u)\rangle\, u'_i u'_j}{\vartheta^2}$.

Since the quantities $\langle N(u),\mathbf{r}_{ij}(u)\rangle$ are determined by the parameterization of the hypersurface, the functions $u'_1,\dots,u'_{n-1}$ are the components of the speed vector γ' of the curve with respect to the basis $\mathbf{r}_1,\dots,\mathbf{r}_{n-1}$ of the tangent space, ϑ is the length of the speed vector γ', $k(\gamma')$ depends only on the speed vector γ' of the curve (that justifies the notation $k(\gamma')$).

Definition. Let $\mathbf{v}$ be an arbitrary tangent vector of the regular parameterized hypersurface $\mathbf{r}:\Omega\longrightarrow \mathbb{R}^n$ at $\mathbf{r}(u)$, The intersection curve of the hypersurface and the plane through $\mathbf{r}(u)$ spanned by direction vectors $N(u)$ and $\mathbf{v}$ is called the normal section of the hypersurface in the direction $\mathbf{v}$. Giving the cutting normal plane an orientation by the ordered basis $(\mathbf{v}, N(u))$, we may consider the signed curvature of the normal section, which will be called the normal curvature of the hypersurface in the direction $\mathbf{v}$ and will be denoted by $k(\mathbf{v})$.

Applying the above general formulas for normal sections one may see easily that the normal curvature of a parameterized hypersurface in the direction $\mathbf{v} = v_1\, \mathbf{r}_1(u) +\dots+ v_{n-1}\, \mathbf{r}_{n-1}(u)$ is just

$$k(\mathbf{v}) = \frac{\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \langle N(u),\mathbf{r}_{ij}(u)\rangle\, v_i v_j}{\vartheta^2} ,$$

where $\vartheta = \|\mathbf{v}\|$. Since $k(\lambda\mathbf{v}) = k(\mathbf{v})$ for any $\lambda \neq 0$, the normal curvature depends only on the straight line of $\mathbf{v}$.

Returning to the curve γ we see that $k(\gamma')$ is determined by the tangent line of γ at the given point which is the intersection of the osculating plane of γ and the tangent space of the hypersurface.

Since the osculating plane and the tangent line determines the second Frenet vector $\mathbf{t}_2$ uniquely up to sign, we conclude that the the curvature $\kappa_1 = \left(1/\langle N(u),\mathbf{t}_2\rangle\right) k(\gamma')$ of the curve is determined by the osculating plane up to sign and since the curvature κ_1 is positive, both $\mathbf{t}_2$ and κ_1 are determined uniquely (and not only up to sign) by the osculating plane.

To finish this unit with, we formulate an obvious consequence of the formula expressing the curvature of a curve lying on a hypersurface.

Corollary. (Meusnier's theorem) If the osculating plane of a curve γ lying on a hypersurface is not contained in the tangent space of the hypersurface at a given point $\gamma(t) = \mathbf{r}(u(t))$, then the curvature of the curve and the normal curvature of the surface in the direction $\gamma'(t)$ are related to one another by the equation $\kappa_1(t) = \frac{1}{\cos\alpha} k(\gamma'(t))$, where α is the angle between the normal vector $N(u(t))$ of the hypersurface and the second Frenet vector $\mathbf{t}_2(t)$ of the curve.

Further Exercises

4-1. Given a unit speed curve of general type in $\mathbb{R}^3$ with distinguished Frenet frame $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3$, find a vector field ω along the curve such that $\dot{\mathbf{t}}_i = \omega \times \mathbf{t}_i$ holds for i=1,2,3. (ω is called the Darboux vector field of the curve.)

4-2. Suppose that the osculating planes of a curve of general type in $\mathbb{R}^3$ have a point in common. Show that the curve is a plane curve.

4-3. Suppose that the normal planes of a regular curve in $\mathbb{R}^3$ go through a fixed point O. Show that the curve lies on a sphere centered at O.

4-4. Let γ be a curve of general type in $\mathbb{R}^n$, $\mathbf{t}_1, \dots \mathbf{t}_n$ its distinguished Frenet frame, $0 \leqslant k \leqslant n$. By the definition of the distinguished Frenet frame, the k-th derivative of γ can be expressed as a linear combination of $\mathbf{t}_1, \dots, \mathbf{t}_k$ as

$$\gamma^{(k)} = c_1\,\mathbf{t}_1 + \dots + c_k\,\mathbf{t}_k\ ,$$

where $c_1, \dots, c_k$ are suitable functions. Show that

$$c_k = |\gamma'|^k\,\kappa_1\,\kappa_2 \dots \kappa_{k-1}\ .$$

4-5. Compute the curvatures of the "moment curve" $\gamma(t) = (t, t^2, \dots, t^n)$ at $t = 0$. (Hint: Use previous exercise.)

Unit 5. Hypersurfaces

Vector fields along hypersurfaces, tangential vector fields, derivations of vector fields with respect to a tangent direction, the Weingarten map, bilinear forms, the first and second fundamental forms of a hypersurface, principal directions and principal curvatures, mean curvature and the Gaussian curvature, Euler's formula.

Definition. Let $\mathbf{r}:\Omega \longrightarrow \mathbb{R}^n$ be a parameterized hypersurface. A vector field along the hypersurface is a mapping $X:\Omega \longrightarrow T_*\mathbb{R}^n$ from the domain of parameters into the tangent bundle of $\mathbb{R}^n$ such that $X(u) \in T_{\mathbf{r}(u)}\mathbb{R}^n$ for any $u \in \Omega$. X is a tangential vector field, if X(u) is tangent to the hypersurface at $\mathbf{r}(u)$.

Since X(u) has the form $(\mathbf{r}(u),\tilde{X}(u))$, where $\tilde{X}:\Omega \longrightarrow \mathbb{R}^n$, there is a one to one correspondence between vector fields along a parameterized hypersurface and smooth mappings of the domain of parameters into $\mathbb{R}^n$. Roughly speaking, if we are given a smooth mapping of the parameter domain into $\mathbb{R}^n$, we may think of it as a vector field along the hypersurface though formally it is not a vector field. In this way, the mappings $\mathbf{r}_1,\ldots,\mathbf{r}_{n-1}$ and N should be thought of as vector fields along the hypersurface, the first n-1 of which are tangential.

Given a vector field along a hypersurface, we would like to express the speed of change of the vector field vectors as we move along the surface, in terms of the speed of our motion. This is achieved by the following.

Definition. Let $\mathbf{r}:\Omega \longrightarrow \mathbb{R}^n$ be a parameterized hypersurface, $X:\Omega \longrightarrow \mathbb{R}^n$ be a vector field along it, $u_0 \in \Omega$, $\mathbf{v}$ a tangent vector of the hypersurface at $\mathbf{r}(u_0)$. We define the derivative $\partial_{\mathbf{v}}X$ of the vector field X in the direction $\mathbf{v}$ as $\partial_{\mathbf{v}}X = (X\circ u)'(0)$, where $u: [-1,1] \longrightarrow \Omega$ is a curve in the parameter domain such that $u(0) = u_0$ and $(\mathbf{r}\circ u)'(0) = \mathbf{v}$.

Since by the chain rule

$$(X\circ u)'(0) = \sum_{i=1}^{n-1} u_i'(0)\; X_i(u(0)) \;,$$

where $(u_1,\ldots,u_{n-1})$ are the components of u, $X_1,\ldots,X_{n-1}$ are the partial derivatives of X, and by

$$\mathbf{v} = (\mathbf{r}\circ u)'(0) = \sum_{i=1}^{n-1} u_i'(0)\; \mathbf{r}_i(u(0))$$

the numbers $u'_1(0),\dots,u'_{n-1}(0)$ are the components of $\mathbf{v}$ in the basis $\mathbf{r}_1(u_0),\dots,\mathbf{r}_{n-1}(u_0)$ of the tangent space at $\mathbf{r}(u_0)$, we have the following formula

$$\partial_{\mathbf{v}} X = \sum_{i=1}^{n-1} v_i \ X_i(u_0),$$

where $v_1,\dots,v_{n-1}$ are the components of the vector $\mathbf{v}$ in the basis $\mathbf{r}_1(u_0),\dots,$ $\mathbf{r}_{n-1}(u_0)$. This formula shows that the definition of $\partial_{\mathbf{v}}X$ is correct, i.e. independent of the choice of the curve $u(t)$.

We shall consider the local behavior of curvature on a hypersurface. The way in which a hypersurface curves around in $\mathbb{R}^n$ is closely related to the way the normal direction changes as we move from point to point.

Lemma. The derivative $\partial_{\mathbf{v}}N$ of the normal direction on a hypersurface with respect to a tangent vector $\mathbf{v}$ at $p = \mathbf{r}(u)$ is tangent to the hypersurface at $\mathbf{r}(u)$.

Proof. We need to show that $\partial_Y N$ is orthogonal to $N(p)$. Indeed, differentiating the relation $1 \equiv \langle N , N \rangle$, we get

$$0 = \langle \partial_{\mathbf{v}}N , N \rangle + \langle N , \partial_{\mathbf{v}}N \rangle = 2 \langle \partial_{\mathbf{v}}N , N \rangle. \blacksquare$$

Definition. Let us denote by M the parameterized hypersurface $\mathbf{r}:\Omega \longrightarrow \mathbb{R}^n$ and by T_pM the linear space of its tangent vectors at $p = \mathbf{r}(u_0)$. The linear map

$$L_p : T_pM \longrightarrow T_pM$$

defined for a fixed $p \in M$ by

$$L_p(\mathbf{v}) = - \partial_{\mathbf{v}} N$$

is called the Weingarten map or shape operator of M at p.

Before going on with the study of hypersurfaces, let us recall some definitions from linear algebra.

Definition. Let V be a vector space. A bilinear function or form on V is a mapping

$$B : V \times V \longrightarrow \mathbb{R}$$

satisfying the identities

$$B(\alpha\mathbf{x}_1+\beta\mathbf{x}_2,\mathbf{y}) = \alpha B(\mathbf{x}_1,\mathbf{y}) + \beta B(\mathbf{x}_2,\mathbf{y}),$$
$$B(\mathbf{x},\alpha\mathbf{y}_1+\beta\mathbf{y}_2) = \alpha B(\mathbf{x},\mathbf{y}_1) + \beta B(\mathbf{x},\mathbf{y}_2)$$

for any $\mathbf{x},\mathbf{x}_1,\mathbf{x}_2,\mathbf{y},\mathbf{y}_1,\mathbf{y}_2 \in V$, $\alpha,\beta \in \mathbb{R}$.

B is said to be symmetric if

$$B(\mathbf{x},\mathbf{y}) = B(\mathbf{y},\mathbf{x}) \qquad \text{for all } \mathbf{x}\ \mathbf{y} \in V.$$

A symmetric bilinear function is positive definite if $B(\mathbf{x},\mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{0}$.

A vector space equipped with a positive definite symmetric bilinear form is a Euclidean vector space.

For example, $\mathbb{R}^n$ with the usual dot product on it is a Euclidean vector space.

If $\mathbf{x}_1,\dots,\mathbf{x}_n$ is a basis of the vector space V and B is a bilinear function on V then the nxn matrix (b_{ij}) with entries $b_{ij} = B(\mathbf{x}_i,\mathbf{x}_j)$ is called the <u>matrix representation of</u> B <u>with respect to the basis</u> $\mathbf{x}_1,\dots,\mathbf{x}_n$. Fixing the basis we get a one to one correspondence between bilinear functions and nxn matrices. A bilinear form is symmetric if and only if its matrix representation with respect to a basis is symmetric.

<u>Definition</u>. The <u>quadratic form of a bilinear function</u> B is the function defined by the equality $Q_B(\mathbf{x}) = B(\mathbf{x},\mathbf{x})$.

Symmetric bilinear functions can be recovered from their quadratic forms with the help of the identity

$$B(\mathbf{x},\mathbf{y}) = \frac{1}{2}\Big(Q_B(\mathbf{x}+\mathbf{y}) - Q_B(\mathbf{x}) - Q_B(\mathbf{y})\Big).$$

Now we return to hypersurfaces. We define two bilinear forms on each tangent space of the hypersurface

<u>Definition</u>. Let M be a parameterized hypersurface $\mathbf{r}:\Omega\longrightarrow\mathbb{R}^n$, $u_0\in\Omega$, T_pM the linear space of tangent vectors of M at $p = \mathbf{r}(u)$, $L_p:T_pM\longrightarrow T_pM$ the Weingarten map. The <u>first fundamental form</u> of the hypersurface is the bilinear function I_p on T_pM obtained by restriction of the dot product onto T_pM

$$I_p(\mathbf{v},\mathbf{w}) = \langle\, \mathbf{v}\ ,\ \mathbf{w}\,\rangle \qquad \text{for } \mathbf{v},\mathbf{w}\in T_pM.$$

The <u>second fundamental form</u> of the hypersurface is the bilinear function II_p on T_pM defined by the equality

$$II_p(\mathbf{v},\mathbf{w}) = \langle\, L_p\mathbf{v}\ ,\ \mathbf{w}\,\rangle \qquad \text{for } \mathbf{v},\mathbf{w}\in T_pM.$$

The first fundamental form is obviously a positive definite symmetric bilinear function on the tangent space. Its matrix representation with respect to the basis $\mathbf{r}_1(u_0),\dots,\mathbf{r}_{n-1}(u_0)$ has entries $\langle\, \mathbf{r}_i(u_0)\ ,\ \mathbf{r}_j(u_0)\,\rangle$.

An important property of the Weingarten map and the second fundamental form is stated in the following theorem.

<u>Theorem</u>. The second fundamental form is symmetric, i.e.

$$\langle\, L_p\mathbf{v}\ ,\ \mathbf{w}\,\rangle = \langle\, \mathbf{v}\ ,\ L_p\mathbf{w}\,\rangle \qquad \text{for } \mathbf{v},\mathbf{w}\in T_pM,$$

or in other words, the Weingarten map is self-adjoint (with respect to the first fundamental form).

<u>Proof</u>. It is enough to prove that the matrix of the second fundamental form with respect to the basis $\mathbf{r}_1(u_0),\dots,\mathbf{r}_{n-1}(u_0)$ is symmetric.

<u>Lemma</u>. $II_p(\mathbf{r}_i(u_0),\mathbf{r}_j(u_0)) = \langle\, \mathbf{r}_{ij}(u_0)\ ,\ N(u_0)\,\rangle$.

<u>Proof of Lemma</u>. We know that the normal vector field N is perpendicular to

any tangential vector field, thus

$$\langle N , \mathbf{r}_j \rangle \equiv 0.$$

Differentiating this identity with respect to the i-th parameter we get

$$\langle \partial_{\mathbf{r}_i} N , \mathbf{r}_j \rangle + \langle N , \mathbf{r}_{ji} \rangle \equiv 0,$$

from which

$$\langle N , \mathbf{r}_{ji} \rangle \equiv \langle - \partial_{\mathbf{r}_i} N , \mathbf{r}_j \rangle = \langle L_{\mathbf{r}} \mathbf{r}_i , \mathbf{r}_j \rangle. \blacksquare$$

Since by Young's theorem $\mathbf{r}_{ij} = \mathbf{r}_{ji}$, the lemma shows that the matrix of the second fundamental form is symmetric. $\blacksquare$

Comparing the identity proved in the lemma with the formula expressing the normal curvature of the hypersurface in a tangent direction $\mathbf{v}$ we see that the normal curvature is the quotient of the quadratic forms of the second and first fundamental forms

$$k(\mathbf{v}) = \frac{\sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \langle N(u_0), \mathbf{r}_{ij}(u_0) \rangle v_i v_j}{\upsilon^2} = \frac{II_p(\mathbf{v},\mathbf{v})}{I_p(\mathbf{v},\mathbf{v})},$$

where $\mathbf{v} = \sum_{i=1}^{n-1} v_i \mathbf{r}_i(u_0)$ is a tangent vector of the hypersurface at $p = \mathbf{r}(u_0)$.
The expression

$$\boxed{k(\mathbf{v}) = \frac{II_p(\mathbf{v},\mathbf{v})}{I_p(\mathbf{v},\mathbf{v})}}$$

gives rise to a linear algebraic investigation of the normal curvature.

It is natural to ask at which directions the normal curvature attains its extrema. Since $k(\lambda\mathbf{v}) = k(\mathbf{v})$ for any $\lambda \neq 0$, it is enough to consider this question for the restriction of k onto the unit sphere S in the tangent space. The unit sphere of a Euclidean space is a compact (=closed and bounded) subset, thus by Weierstrass theorem, any continuous function defined on it attains its maximum and minimum.

Definition. Let f be a differentiable function defined on the unit sphere S of a Euclidean vector space. We say that the vector $\mathbf{v} \in S$ is a critical point of f if for any curve $\gamma : [-1,1] \longrightarrow S$ on the sphere such that $\gamma(0) = \mathbf{v}$ the derivative of the composite function $f \circ \gamma$ vanishes at 0.

Clearly, local minimum and maximum points of a function are its critical points, but the converse is not true. The following proposition gives a characterization of critical points for the restriction of the normal curvature onto the unit sphere of the tangent space.

Proposition. Let V be a finite dimensional vector space with a positive definite symmetric bilinear function <,> and let $L:V\longrightarrow V$ be a self-adjoint linear transformation on V. Let $S = \{\mathbf{x} \in V : \langle \mathbf{x},\mathbf{x}\rangle = 1\}$ and define $f:S\longrightarrow \mathbb{R}$ by $f(\mathbf{x})=\langle L\mathbf{x},\mathbf{x}\rangle$. Then $\mathbf{v} \in S$ is a critical point of f if and only if $\mathbf{v}$ is an eigenvector of L.

Proof. For any curve $\gamma:[-1,1]\longrightarrow S$ such that $\gamma(0) = \mathbf{v}$, we have

$$\frac{d}{dt}\langle L(\gamma(t)),\gamma(t)\rangle\Big|_{t=0} = \langle L(\gamma'(0)),\gamma(0)\rangle+\langle L(\gamma(0)),\gamma'(0)\rangle$$
$$= \langle L\gamma'(0),\mathbf{v}\rangle+\langle L\mathbf{v},\gamma'(0)\rangle = 2\langle L\mathbf{v},\gamma'(0)\rangle.$$

This means that $\mathbf{v}$ is a critical point of f if and only if $L\mathbf{v}$ is orthogonal to every vectors of the form $\gamma'(0)$. Since the speed vectors $\gamma'(0)$ of spherical curves through $\mathbf{v} = \gamma(0)$ range over the tangent space of the sphere S, $\mathbf{v}$ is a critical point of f if and only if $L\mathbf{v}$ is orthogonal to the tangent space of S at $\mathbf{v}$ however, since the normal vector of this tangent space is $\mathbf{v}$, the latter condition is satisfied if and only if $L\mathbf{v}$ is a scalar multiple of $\mathbf{v}$, i.e. $\mathbf{v}$ is an eigenvector of L. ▮

As an application of the proposition, let us prove the following theorem of linear algebra.

Theorem. Let V be a finite dimensional Euclidean vector space and let $L : V \longrightarrow V$ be a self-adjoint linear transformation on V. Then there exists an orthonormal basis of V consisting of eigenvectors of L.

Proof. By induction on the dimension n of V. For $n = 1$ the theorem is trivial. Assume that it is true for $n = k$. Suppose $n = k + 1$. By the proposition, there exists a unit vector $\mathbf{v}_1$ in V which is an eigenvector of L. Let $W = \mathbf{v}_1^{\perp} = \{\mathbf{w} \in V : \mathbf{v}_1 \perp \mathbf{w}\}$. Then $L(W) \subseteq W$ since we have

$$\langle L\mathbf{w},\mathbf{v}_1\rangle = \langle \mathbf{w},L\mathbf{v}_1\rangle = \langle \mathbf{w},\lambda_1\mathbf{v}_1\rangle = \lambda_1\langle \mathbf{w},\mathbf{v}_1\rangle = 0$$

for any $\mathbf{w} \in W$, where λ_1 is the eigenvalue belonging to $\mathbf{v}_1$. Clearly $L|_W$ is self-adjoint. Since $\dim(W) = \dim(V) - 1 = k$, the induction assumption implies that there exists an orthonormal basis $(\mathbf{v}_2,\ldots,\mathbf{v}_n)$ for W consisting of eigenvectors of $L|_W$. But each eigenvector of $L|_W$ is an eigenvector of L, so $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ is an orthonormal basis for V consisting of eigenvectors of L. ▮

Definition. For a hypersurface M in $\mathbb{R}^n$ parameterized by $\mathbf{r}$, $\mathbf{r}(u_0) = p \in M$, the eigenvalues $\kappa_1(p),\ldots,\kappa_{n-1}(p)$ of the Weingarten map $L_p:T_pM\longrightarrow T_pM$ are called the principal curvatures of M at p, the unit eigenvectors of L_p are called principal curvature directions.

If the principal curvatures are ordered so that $\kappa_1(p)\leqslant\kappa_2(p)\leqslant\ldots\leqslant\kappa_{n-1}(p)$, the discussion above shows that $\kappa_{n-1}(p)$ is the maximal, $\kappa_1(p)$ is the minimal value of the normal curvature $k(\mathbf{v})$.

Theorem. (Euler's formula) Let $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$ be an orthonormal basis of T_pM consisting of principal curvature directions, $\kappa_1(p),\dots,\kappa_{n-1}(p)$ be the corresponding principal curvatures. Then the normal curvature $k(\mathbf{v})$ in the direction $\mathbf{v} \in T_pM$, $\|\mathbf{v}\| = 1$, is given by

$$k(Y) = \sum_{i=1}^{n-1} k_i(p) \langle Y,Y_i\rangle^2 = \sum_{i=1}^{n-1} k_i(p) \cos^2(\theta_i),$$

where $\theta_i = \arccos(\langle \mathbf{v},\mathbf{v}_i\rangle)$ is the angle between $\mathbf{v}$ and $\mathbf{v}_i$.

Proof. Since $(\mathbf{v}_1,\dots,\mathbf{v}_n)$ is an orthonormal basis, the vector $\mathbf{v}$ can be expressed as

$$\mathbf{v} = \sum_{i=1}^{n-1} \langle \mathbf{v},\mathbf{v}_i\rangle \mathbf{v}_i.$$

Making use of this formula, we obtain

$$k(\mathbf{v}) = \langle L_p(\mathbf{v}),\mathbf{v}\rangle = \langle L_p\left(\sum_{i=1}^{n-1} \langle \mathbf{v},\mathbf{v}_i\rangle \mathbf{v}_i\right), \sum_{i=1}^{n-1} \langle \mathbf{v},\mathbf{v}_i\rangle \mathbf{v}_i\rangle =$$

$$= \langle \sum_{i=1}^{n-1} \langle \mathbf{v},\mathbf{v}_i\rangle \kappa_i(p) \mathbf{v}_i, \sum_{i=1}^{n-1} \langle \mathbf{v},\mathbf{v}_i\rangle \mathbf{v}_i\rangle = \sum_{i=1}^{n-1} \kappa_i(p) \langle \mathbf{v},\mathbf{v}_i\rangle^2. \blacksquare$$

The determinant and trace of the Weingarten map, that is the product and sum of the principal curvatures are of particular importance in differential geometry.

Definition. For M a hypersurface, $p \in M$, the determinant $K(p)$ of the Weingarten map L_p is called the Gaussian or Gauss-Kronecker curvature of M at p, $H(p) = 1/(n-1)\,\mathrm{trace}\,(L_p)$ is called the mean curvature.

When we compute the principal curvatures and directions of a hypersurface at a point we generally work with a matrix representation of the Weingarten map. Recall that if V is a linear space with basis $\mathbf{x}_1,\dots,\mathbf{x}_n$ and $L:V\longrightarrow V$ is a linear mapping then the matrix representation of V with respect to the basis $\mathbf{x}_1,\dots,\mathbf{x}_n$ is the nxn matrix (ℓ_{ij}) for which

$$L(\mathbf{x}_i) = \sum_{j=1}^{n} \ell_{ij}\mathbf{x}_j \qquad i=1,2,\dots,n.$$

When we have a deal with a parameterized hypersurface $\mathbf{r}:\Omega\longrightarrow \mathbb{R}^n$, it is natural to take the basis $\mathbf{r}_1(u),\dots,\mathbf{r}_{n-1}(u)$ of the tangent space at $\mathbf{r}(u)$. Let us denote by $\mathcal{G} = (g_{ij})$, $\mathcal{B} = (b_{ij})$ and $\mathcal{L} = (\ell_{ij})$ the matrix representations of the first and second fundamental forms and the Weingarten map respectively, with respect to this basis (g_{ij}, b_{ij} and ℓ_{ij} are functions on the parameter domain). Components of $\mathcal{G}$ and $\mathcal{B}$ can be calculated according to the equations

$$g_{ij} = \langle \mathbf{r}_i , \mathbf{r}_j \rangle,$$
$$b_{ij} = \langle N , \mathbf{r}_{ij} \rangle \quad \text{(cf. Lemma above).}$$

The relationship between the matrices $\mathcal{G}$, $\mathcal{B}$, and $\mathcal{L}$ follows from the following equalities

$$b_{ij} = \langle L_{\mathbf{r}}\mathbf{r}_i , \mathbf{r}_j \rangle = \langle \sum_{k=1}^{n-1} \ell_{ik}\mathbf{r}_k , \mathbf{r}_j \rangle = \sum_{k=1}^{n-1} \ell_{ik} \langle \mathbf{r}_k , \mathbf{r}_j \rangle =$$

$$= \sum_{k=1}^{n-1} \ell_{ik} \, g_{kj}$$

expressing that $\mathcal{B} = \mathcal{L}\,\mathcal{G}$. $\mathcal{G}$ is the matrix of a positive definite bilinear function, hence it is invertible (its determinant is positive). Multiplying the equation $\mathcal{B} = \mathcal{L}\,\mathcal{G}$ with the inverse of $\mathcal{G}$ we get the expression

$$\boxed{\mathcal{L} = \mathcal{B}\,\mathcal{G}^{-1}}$$

for the matrix of the Weingarten operator.

Corollary. The Gaussian curvature of a hypersurface is equal to

$$\boxed{K = \frac{\det \mathcal{B}}{\det \mathcal{G}}}.$$

Recall from linear algebra that in order to determine the eigenvalues of a linear mapping with matrix representation $\mathcal{L}$ one has to find the roots of the characteristic polynomial $p_{\mathcal{L}}(\lambda) = \det(\mathcal{L} - \lambda I)$, where I denotes the identity matrix.

Having determined the eigenvalues of the linear mapping, components of eigenvectors with respect to the fixed basis are obtained as non-zero solutions of the linear system of equations $\mathcal{L}\,\mathbf{v} = \lambda\,\mathbf{v}$ where λ is a nonzero eigenvalue of $\mathcal{L}$.

Further Exercices

5-1. Determine the Weingarten map for a sphere of radius r at one of its points.

5-2. Find the normal curvature $k(\mathbf{v})$ for each tangent direction $\mathbf{v}$, the principal curvatures and the principal curvature directions, and compute the Gaussian and mean curvatures of the following surfaces at the given point p.

$(x_1^2/a^2)+(x_2^2/b^2)+(x_3^2/c^2) = 1$, $p = (a,0,0)$ (ellipsoid);

$(x_1^2/a^2)+(x_2^2/b^2)-(x_3^2/c^2) = 1$, $p = (a,0,0)$ (one-sheeted hyperboloid);

$x_1^2 + \left(\sqrt{x_2^2+x_3^2} - 2\right)^2 = 1$ $p = (0,3,0)$ or $p = (0,1,0)$ (torus).

5-3. Suppose that the principal curvatures of a parameterized surface in $\mathbb{R}^3$ vanish. Show that the surface is a part of a plane.

5-4. Find the Gaussian curvature $K:M\longrightarrow \mathbb{R}$ for the following surfaces

$x_1^2+x_2^2-x_3^2 = 0$, $x_3 > 0$ (cone)

$(x_1^2/a^2)+(x_2^2/b^2)-(x_3^2/c^2) = 1$ (hyperboloid);

$(x_1^2/a^2)+(x_2^2/b^2)-x_3 = 0$ (elliptic paraboloid);

$(x_1^2/a^2)-(x_2^2/b^2)-x_3 = 0$ (hyperbolic paraboloid).

5-5. Let M be a (hyper)surface in $\mathbb{R}^3$, $p \in M$. Show that for each $\mathbf{v},\mathbf{w} \in T_pM$,

$$L_p(\mathbf{v}) \times L_p(\mathbf{w}) = K(p)\ \mathbf{v} \times \mathbf{w}.$$

UNIT 6. SURFACES IN THE 3-DIMENSIONAL SPACE

Umbilical, spherical and planar points, surfaces consisting of umbilics; surfaces of revolution, Beltrami's pseudosphere; lines of curvature, parameterizations for which coordinate lines are lines of curvature, Dupin's theorem, confocal second order surfaces; ruled and developable surfaces: equivalent definitions, basic examples, relations to surfaces with K=0, structure theorem.

A regular parameterized surface $\mathbf{r}:\Omega \longrightarrow \mathbb{R}^3$ (Ω is an open subset of the plane) has two principal curvatures $\kappa_1(u,v)$ and $\kappa_2(u,v)$ at each point $p = \mathbf{r}(u,v)$ of the surface. If $\kappa_1(u,v) \leqslant \kappa_2(u,v)$ then $\kappa_1(u,v)$ is the minimum of normal curvatures in different directions at p, while $\kappa_2(u,v)$ is the maximum of them. If $\kappa_1(u,v) < \kappa_2(u,v)$ then the principal directions corresponding to $\kappa_1(u,v)$ and $\kappa_2(u,v)$ are uniquely defined, however if $\kappa_1(u,v) = \kappa_2(u,v)$ then the normal curvature is constant in all directions and every direction is principal.

Definition. A point $p = \mathbf{r}(u,v)$ of a surface is called a umbilical point or umbilic if the principal curvatures at p are equal. A umbilical point p is said to be spherical if $\kappa_1(u,v) = \kappa_2(u,v) \neq 0$, and planar if $\kappa_1(u,v) = \kappa_2(u,v) = 0$.

The following theorem gives a characterization of those surfaces which have only umbilical points.

Theorem. A connected regular surface all points of which are umbilical is contained in a plane or sphere.

Proof. First we show that the principal curvature function $\kappa_1 = \kappa_2 = \kappa$ is constant along the surface. Fixing a parameterization $\mathbf{r}$, we have

$$N_u = -\kappa\, \mathbf{r}_u \quad \text{and} \quad N_v = -\kappa\, \mathbf{r}_v,$$

since $\mathbf{r}_u$ and $\mathbf{r}_v$ are principal directions as any tangent vector is. Differentiating the first equation with respect to v, the second with respect to u, we get

$$N_{uv} = -\kappa_v\, \mathbf{r}_u - \kappa\, \mathbf{r}_{uv} \quad \text{and} \quad N_{uv} = -\kappa_u\, \mathbf{r}_v - \kappa\, \mathbf{r}_{uv},$$

from which $\kappa_v\, \mathbf{r}_u = \kappa_u\, \mathbf{r}_v$. Since $\mathbf{r}_u$ and $\mathbf{r}_v$ are linearly independent, the last equation can hold only if $\kappa_u = \kappa_v = 0$, i.e. if κ is constant.

1^{st} case: $\kappa \equiv 0$. In this case $N_u = N_v = 0$, therefore the normal vector is constant along the surface. The derivative of the function $\langle N, \mathbf{r} \rangle$ with respect to u and v is $\langle N,\mathbf{r}_u\rangle = \langle N,\mathbf{r}_v\rangle = 0$ because N is perpendicular to the tangent vectors $\mathbf{r}_u,\mathbf{r}_v$, hence $\langle N,\mathbf{r}\rangle$ is constant and the surface is contained in a plane with equation $\langle N,\mathbf{x}\rangle = \text{const}$.

2nd case: $\kappa \neq 0$. We claim that in this case the surface is contained in a sphere. The facts we have so far suggests that if the claim is true then the center of the sphere should be $\mathbf{r} + (1/\kappa)\, N$. Setting $\mathbf{p} = \mathbf{r} + (1/\kappa)\, N$, we have to make sure first that $\mathbf{p}$ does not depend on u and v. Indeed, differentiating with respect to u results

$$\mathbf{p}_u = \mathbf{r}_u + (1/\kappa)\, N_u = \mathbf{r}_u - (1/\kappa)\,\kappa\, \mathbf{r}_u = 0 \qquad (\kappa \text{ is constant!})$$

and similarly,

$$\mathbf{p}_v = \mathbf{r}_v + (1/\kappa)\, N_v = \mathbf{r}_v - (1/\kappa)\,\kappa\, \mathbf{r}_v = 0.$$

Now to show that the surface lies on a sphere centered at $\mathbf{p}$ we have to prove that the function $\|\mathbf{r}-\mathbf{p}\|$ is constant. This follows from

$$\frac{\partial}{\partial u}\|\mathbf{r}-\mathbf{p}\|^2 = 2\langle \mathbf{r}_u, \mathbf{r}-\mathbf{p}\rangle = 2\langle \mathbf{r}_u, (1/\kappa)N\rangle = 0,$$

and

$$\frac{\partial}{\partial v}\|\mathbf{r}-\mathbf{p}\|^2 = 2\langle \mathbf{r}_v, \mathbf{r}-\mathbf{p}\rangle = 2\langle \mathbf{r}_v, (1/\kappa)N\rangle = 0.$$

The theorem is proved. ▮

The next example shows how to compute the principal curvatures and directions for a surface of revolution. We consider a positive function $f:[a,b]\longrightarrow \mathbb{R}_+$ and the surface of revolution generated by rotation of its graph about x-axis. This surface can be parameterized by the mapping

$$\mathbf{r}(u,v) = \big(\, u,\ f(u)\cos v,\ f(u)\sin v\,\big).$$

The tangent vectors $\mathbf{r}_u$ and $\mathbf{r}_v$ are obtained by partial differentiation

$$\mathbf{r}_u(u,v) = \big(\, 1,\ f'(u)\cos v,\ f'(u)\sin v\,\big),$$
$$\mathbf{r}_v(u,v) = \big(\, 0, -f(u)\sin v,\ f(u)\cos v\,\big).$$

The matrix of the first fundamental form with respect to the basis $\mathbf{r}_u,\mathbf{r}_v$ is

$$\mathcal{G} = \begin{pmatrix} \langle \mathbf{r}_u,\mathbf{r}_u\rangle & \langle \mathbf{r}_u,\mathbf{r}_v\rangle \\ \langle \mathbf{r}_v,\mathbf{r}_u\rangle & \langle \mathbf{r}_v,\mathbf{r}_v\rangle \end{pmatrix} = \begin{pmatrix} 1 + f'^2(u) & 0 \\ 0 & f^2(u) \end{pmatrix}.$$

To obtain the matrix of the second fundamental form we need the normal vector field and the second order partial derivatives of $\mathbf{r}$.

$$\mathbf{r}_u\times \mathbf{r}_v = \det\begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & f'(u)\cos v & f'(u)\sin v \\ 0 & -f(u)\sin v & f(u)\cos v \end{pmatrix} =$$

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

$$N = \frac{\mathbf{r}_u\times \mathbf{r}_v}{\|\mathbf{r}_u\times \mathbf{r}_v\|} = \frac{1}{\sqrt{1 + f'^2(u)}}\big(f'(u),\ -\cos v,\ -\sin v\big)$$

$$\mathbf{r}_{uu}(u,v) = \big(0,\ f''(u)\cos v,\ f''(u)\sin v \big),$$
$$\mathbf{r}_{uv}(u,v) = \big(0, - f'(u)\sin v,\ f'(u)\cos v \big),$$
$$\mathbf{r}_{vv}(u,v) = \big(0,\ - f(u)\cos v,\ - f(u)\sin v \big).$$

$$\mathcal{B} = \begin{pmatrix} \langle N,\mathbf{r}_{uu}\rangle & \langle N,\mathbf{r}_{uv}\rangle \\ \langle N,\mathbf{r}_{vu}\rangle & \langle N,\mathbf{r}_{vv}\rangle \end{pmatrix} = \frac{1}{\sqrt{1+f'^2(u)}} \begin{pmatrix} -f''(u) & 0 \\ 0 & f(u) \end{pmatrix}.$$

The matrix of the Weingarten map with respect to the basis $\mathbf{r}_u, \mathbf{r}_v$ is

$$\mathcal{L} = \mathcal{B}\,\mathcal{G}^{-1} = \begin{pmatrix} -f''(u)\big(1+f'^2(u)\big)^{-3/2} & 0 \\ 0 & f^{-1}(u)\big(1+f'^2(u)\big)^{-1/2} \end{pmatrix}.$$

As we see, the matrix of the Weingarten map is diagonal, consequently $\mathbf{r}_u, \mathbf{r}_v$ are eigenvectors the diagonal elements of $\mathcal{L}$ are eigenvalues of the Weingarten map. Thus the principal curvatures of the surface are

$$\kappa_1(u,v) = -\frac{f''(u)}{\big(1+f'^2(u)\big)^{3/2}} \qquad \kappa_2(u,v) = \frac{1}{f(u)\big(1+f'^2(u)\big)^{1/2}}\ .$$

We could have obtained this result in a more geometrical way. For any point p on the surface, the plane through p and the x-axis is a symmetry plane of the surface. Thus, reflection of a principal direction of the surface at p is also a principal direction (with the same principal curvature). The principal curvatures at p are either equal and then every direction is principal, or different and then the principal directions are unique. Since a direction is invariant under a reflection in a plane if and only if it is parallel or orthogonal to the plane, we may conclude that $\mathbf{r}_u$ and $\mathbf{r}_v$ are principal directions of the surface. Principal curvatures are the curvatures of the normal sections of the surface in the direction $\mathbf{r}_u, \mathbf{r}_v$.

The normal section of the surface in the direction $\mathbf{r}_u$ is the graph of f rotated about the x-axis (a <u>meridian</u> of the surface). Its curvature can be calculated according to the formula known for plane curves and gives κ_1 up to sign. The difference in sign is due to the fact that the unit normal of the surface and the principal normal of the meridian are opposite to one another.

The plane passing through p perpendicular to the x-axis intersects the surface in a circle the tangent of which at p is $\mathbf{r}_v$. The curvature of this circle is $\frac{1}{f(u)}$. The normal curvature $\kappa_2 = k(\mathbf{r}_v)$ of the surface in the direction $\mathbf{r}_v$ and the curvature of the circle intersection are related to one another by Meusnier's theorem as follows

$$\frac{1}{f(u)} = \frac{1}{\cos\alpha}\,\kappa_2\ ,$$

where α is the angle between the normal of the surface and the principal normal of the circle. As it is easy to see, α is the direction angle of the tangent to the meridian at p, that is, by elementary calculus

$\operatorname{tg} \alpha = f'(u)$, from which $\cos \alpha = \dfrac{1}{\sqrt{1 + f'^2(u)}}$.

Therefore, we get

$$\kappa_2 = \frac{1}{f(u)\big(1 + f'^2(u)\big)^{1/2}}$$

as before. The equation $\dfrac{1}{f(u)} = \dfrac{1}{\cos \alpha}\,\kappa_2$ has the following consequence.

Corollary. The second principal radius of curvature $\dfrac{1}{\kappa_2}$ of a surface of revolution at a given point p is the length of the segment of the normal of the surface between p and the x-axis intercept.

As an application, let us show that the surface of revolution generated by the tractrix has constant -1 Gaussian curvature. For this reason the surface is called pseudosphere. Its intrinsic geometry is locally the same as that of Bolyai's and Lobatchevsky's hyperbolic plane. This local model of hyperbolic geometry was discovered by Beltrami.

The tractrix is defined as the involute of the chain curve $\gamma(t)=(t,\operatorname{ch} t)$ touching the chain curve at (0,1). The length of the chain curve arc between $\gamma(0)$ and $\gamma(t)$ is

$$\int_0^t \sqrt{\|\gamma'(t)\|}\,dt = \int_0^t \sqrt{1 + \operatorname{sh}^2 t}\,dt = \int_0^t \operatorname{ch} t\,dt = \operatorname{sh} t.$$

This way, the tractrix has the parameterization

$$\hat{\gamma}(t) = \gamma(t) - \operatorname{sh} t\,\frac{\gamma'(t)}{\|\gamma'(t)\|} = (t,\operatorname{ch} t) - \operatorname{sh} t\,\frac{(1,\operatorname{sh} t)}{\operatorname{ch} t} = \Big(t - \operatorname{th} t\ ,\ \frac{1}{\operatorname{ch} t}\Big).$$

As we know from the theory of evolutes and involutes, the chain curve is the evolute of the tractrix, the segment $\gamma(t)\hat{\gamma}(t)$ is normal to the tractrix, and its length is the radius of curvature of the tractrix at $\hat{\gamma}(t)$. This implies from one hand that the first principal curvature of the pseudosphere is $\kappa_1 = -\dfrac{1}{\operatorname{sh} t}$. On the other hand, we obtain that the equation of the normal line of the tractrix at $\hat{\gamma}(t)$ is

$$\frac{y - \operatorname{ch} t}{x - t} = \operatorname{sh} t.$$

The x-intercept is

$$\Big(t - \frac{\operatorname{ch} t}{\operatorname{sh} t}\ ,\ 0\Big).$$

According to the general results on surfaces of revolution, the second principal radius of curvature of the pseudosphere is the distance between $\hat{\gamma}(t)$ and $(t - \operatorname{cth} t, 0)$, i.e.

$$\kappa_2^{-1} = \big|\big(\operatorname{cth} t - \operatorname{th} t,\ (\operatorname{ch} t)^{-1}\big| = \left(\left(\frac{\operatorname{ch}^2 t - \operatorname{sh}^2 t}{\operatorname{ch} t \operatorname{sh} t}\right)^2 + \frac{1}{\operatorname{ch}^2 t} \right)^{1/2} =$$

$$= \left(\frac{1}{\mathrm{sh}^2 t\ \mathrm{ch}^2 t} + \frac{1}{\mathrm{ch}^2 t} \right)^{1/2} = \frac{1}{\mathrm{sh}\ t}\ .$$

This shows that $K = \kappa_1 \kappa_2 \equiv -1$.

Definition. A regular curve on a surface is said to be a line of curvature if the tangent vectors of the curve are principal directions.

There are many parameterizations of a hypersurface. In applications we should always try to find a parameterization that makes solving the problem easier. For theoretical purposes, it is good to know the existence of certain parameterizations that have nice properties. In what follows, we study parameterizations, for which coordinate lines are lines of curvature.

Theorem. The coordinate lines of a regular parameterization $\mathbf{r}:\Omega \longrightarrow \mathbb{R}^n$ of a hypersurface are lines of curvature if the matrices of the first and second fundamental forms are diagonal. The converse is also true if the principal curvatures of the hypersurface are different at each point.

Proof. The matrix $\mathcal{L}$ of the Weingarten map is the quotient $\mathcal{B}\ \mathcal{G}^{-1}$ of the matrices of the first and second fundamental forms. If these matrices are diagonal, then so is $\mathcal{L}$. Obviously, the matrix of a linear map with respect to a basis is diagonal if and only the basis consists of eigenvectors of the linear map. In our case, we get that $\mathbf{r}_1,\dots,\mathbf{r}_{n-1}$ form an eigenvector basis for the Weingarten map, i.e. these vectors are principal directions. Since $\mathbf{r}_i$ is the tangent vector of the i-th family of coordinate lines, the coordinate lines are lines of curvature.

Now suppose that the coordinate lines are lines of curvature and that the principal curvatures $\kappa_1,\dots,\kappa_{n-1}$ corresponding to the principal directions $\mathbf{r}_1,\dots,\mathbf{r}_{n-1}$ are different at every point. In this case,

$$\kappa_i \ \langle \mathbf{r}_i, \mathbf{r}_j \rangle = \langle L\mathbf{r}_i, \mathbf{r}_j \rangle = \langle \mathbf{r}_i, L\mathbf{r}_j \rangle = \kappa_j \ \langle \mathbf{r}_i, \mathbf{r}_j \rangle,$$

$$(\kappa_i - \kappa_j)\langle \mathbf{r}_i, \mathbf{r}_j \rangle = 0,$$

and since $(\kappa_i - \kappa_j) \neq 0$ for $i \neq j$,

$$g_{ij} = \langle \mathbf{r}_i, \mathbf{r}_j \rangle = 0 \quad \text{if } i \neq j.$$

Hence, the matrix $\mathcal{G}$ of the first fundamental form is diagonal. The matrix $\mathcal{L}$ of the Weingarten map is also diagonal by our assumption, consequently the matrix $\mathcal{B} = \mathcal{L}\ \mathcal{G}$ of the second fundamental form is diagonal as well. ▮

Theorem. Suppose that a regular parameterized surface in $\mathbb{R}^3$ has no umbilical points. Then every point of the surface has a neighborhood that admits a reparameterization with respect to which coordinate lines are lines of curvature.

Proof. (sketch) The complete proof of the theorem rests upon some results on ordinary differential equations, so we only indicate the geometrical part of the construction of such a parameterization. Since the surface contains no umbilics, principle directions are uniquely defined at each point and one can find two smooth unit vector fields ξ, η along the surface that show in the principal directions at every point.

As it is known, if we are given a tangential vector field ζ and a point p on a surface, then there exists a curve γ_p on the surface such that $\gamma_p(0) = p$ the speed vector $\gamma_p'(t)$ of the curve is just $\zeta(\gamma_p(t))$ for every t from the domain of γ_p. Such curves are called <u>integral curves of the vector field</u> ζ through p. Every integral curve is contained in a maximal one which is unique.

Let us denote by γ^1_* and γ^2_* the integral curves of the vector fields ξ and η through $*$. They are lines of curvature of the surface. Fix a point p and consider the parameterization that assigns to a pair of real numbers (u,v) the intersection point of the curves $\gamma^1_{(\gamma^2(v))}$ and $\gamma^2_{(\gamma^1(u))}$. This parameterization is well defined in a small neighborhood of the origin in $\mathbb{R}^2$. It is smooth and maps onto a neighborhood of p, while the coordinate lines are certain reparameterizations of the integral curves of the vector fields ξ and η i.e. coordinate lines are lines of curvature. $\blacksquare$

<u>Remarks</u>.

The parameterization constructed above has also the property that the coordinate lines through p are parameterized by arc length.

The theorem does not hold for higher dimensions. (Study where the above proof breaks down.)

Now we give a description of lines of curvature on ellipsoids. Our approach, which is based on Dupin's theorem, works for any surfaces of second order. Dupin's theorem claims that if we have three families of surfaces such that the surfaces of any of the families foliate an open domain in $\mathbb{R}^3$, and surfaces from different families intersect one another orthogonally, then the intersection curves of surfaces from different families are lines of curvature. We can obtain families of surfaces in a natural way considering a curvilinear coordinate system on $\mathbb{R}^3$.

<u>Definition</u>. A <u>curvilinear coordinate system</u> on $\mathbb{R}^3$ is a one to one smooth mapping $\mathbf{r}:\Omega \longrightarrow \mathbb{R}^3$ from an open domain of $\mathbb{R}^3$ onto an open domain of $\mathbb{R}^3$ with smooth inverse. Ω is foliated by planes parallel to one of the three coordinate planes in $\mathbb{R}^3$. The images of these planes are the <u>coordinate surfaces</u> of the coordinate system $\mathbf{r}$. There are three families of coordinate

surfaces. Each family foliates the same domain, the image of $\mathbf{r}$. Coordinate surfaces from different families intersect one another in coordinate lines. We say that $\mathbf{r}$ defines a triply orthogonal system of surfaces, if coordinate surfaces from different families intersect one another orthogonally, or equivalently, if $\langle \mathbf{r}_i, \mathbf{r}_j \rangle = 0$ for $i \neq j$.

Theorem. (Dupin's theorem). If the curvilinear coordinate system $\mathbf{r}: \Omega \longrightarrow \mathbb{R}^3$ defines a triply orthogonal system, then the coordinate lines are lines of curvature on the coordinate surfaces.

Proof. We may consider without loss of generality the surface $(u,v) \longmapsto \mathbf{r}(u,v,w_0)$. It is enough to show that the matrices of the first and second fundamental forms of the surface with respect to the given parameterization are diagonal. The matrix of the first fundamental form is diagonal by our assumption $\langle \mathbf{r}_1, \mathbf{r}_2 \rangle = 0$. The nondiagonal element of the matrix of the second fundamental form is $\langle \mathbf{r}_{12}, N \rangle$, where N is the unit normal of the surface. Since $\mathbf{r}_3$ is parallel to N, $\langle \mathbf{r}_{12}, N \rangle = 0$ will follow from $\langle \mathbf{r}_{12}, \mathbf{r}_3 \rangle = 0$. Differentiating the equation $\langle \mathbf{r}_1, \mathbf{r}_3 \rangle = 0$ with respect to the second variable yields

$$\langle \mathbf{r}_{12}, \mathbf{r}_3 \rangle + \langle \mathbf{r}_1, \mathbf{r}_{23} \rangle = 0$$

and similarly,

$$\langle \mathbf{r}_{23}, \mathbf{r}_1 \rangle + \langle \mathbf{r}_2, \mathbf{r}_{31} \rangle = 0$$
$$\langle \mathbf{r}_{31}, \mathbf{r}_2 \rangle + \langle \mathbf{r}_3, \mathbf{r}_{12} \rangle = 0.$$

Solving this system of linear equations for the unknown quantities $\langle \mathbf{r}_{12}, \mathbf{r}_3 \rangle$, $\langle \mathbf{r}_{23}, \mathbf{r}_1 \rangle$, $\langle \mathbf{r}_{31}, \mathbf{r}_2 \rangle$, we see that they are all zero. ∎

The canonical equation of an ellipsoid has the form

$$\frac{x^2}{A} + \frac{y^2}{B} + \frac{z^2}{C} = 1.$$

Suppose A > B > C. We can embed this surface into a triply orthogonal system of second order surfaces as follows. Consider the surface

$$F_\lambda: \quad \frac{x^2}{A+\lambda} + \frac{y^2}{B+\lambda} + \frac{z^2}{C+\lambda} = 1.$$

F_λ is — an ellipsoid for $\lambda > -C$;

— a one sheeted hyperboloid for $-C > \lambda > -B$;

— a two sheeted hyperboloid for $-B > \lambda > -A$.

In accordance with these cases, we get three families of surfaces. Surfaces obtained by such a perturbation of the equation of a second order surface are called confocal second order surfaces.

Proposition. Let $(x,y,z) \in \mathbb{R}^3$ be a point for which $xyz \neq 0$. Then there exist exactly three λ-s, one from each of the intervals $(-C,+\infty)$, $(-B,-C)$,

$(-A,-B)$ such that $(x,y,z) \in F_\lambda$.

Proof. Condition $(x,y,z) \in F_\lambda$ is equivalent to

$$P(\lambda) = (A+\lambda)(B+\lambda)(C+\lambda) - \left(x^2(B+\lambda)(C+\lambda) + y^2(A+\lambda)(C+\lambda) + z^2(A+\lambda)(B+\lambda)\right) = 0.$$

This is an equation of degree three for λ, so the number of solutions is not more than three. To see that all the three solutions are real and located as it is stated, compute P at the nodes.

$$P(-A) = -x^2(B-A)(C-A) < 0$$
$$P(-B) = -y^2(A-B)(C-B) > 0$$
$$P(-C) = -z^2(A-C)(B-C) < 0$$

Furthermore, $F(\lambda) = \lambda^3 + \ldots$ implies $\lim_{\lambda\to\infty} F(\lambda) = +\infty$. Thus by Bolzano's theorem F has at least one root on each of the intervals $(-C,+\infty)$, $(-B,-C)$, $(-A,-B)$. ∎

Proposition. If $(x,y,z) \in F_\lambda \cap F_{\lambda'}$, $xyz \neq 0$, $\lambda \neq \lambda'$, then F_λ intersects $F_{\lambda'}$ orthogonally.

Lemma. If a nonempty subset M of $\mathbb{R}^3$ is defined by an equation

$$M = \{(x,y,z) \in \mathbb{R}^3 : F(x,y,z) = 0\},$$

so that the gradient vector field $\operatorname{grad} F = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right)$ is not zero at points of M, then every point of M has a neighborhood (in M) which is the image of a regular parameterized surface. In this case the tangent plane of M at $p \in M$ is orthogonal to grad F (p).

The first part of the lemma is a direct application of the implicit function theorem, we omit details. Suppose that M admits a regular parameterization $\mathbf{r}$ around $p \in M$. Then $F\circ\mathbf{r} \equiv 0$. Differentiating with respect to the i-th variable (i=1,2) using the chain rule we obtain

$$0 = \frac{\partial F\circ\mathbf{r}}{\partial u_i} = \langle (\operatorname{grad} F)\circ\mathbf{r}, \mathbf{r}_i \rangle,$$

hence grad F (p) is orthogonal to the tangent vectors $\mathbf{r}_1, \mathbf{r}_2$ that span the tangent space.

Proof. (of proposition) We need to show

$$\operatorname{grad} F_\lambda(x,y,z) \perp \operatorname{grad} F_{\lambda'}(x,y,z),$$

or equivalently,

$$\frac{x^2}{(A+\lambda)(A+\lambda')} + \frac{y^2}{(B+\lambda)(B+\lambda')} + \frac{z^2}{(C+\lambda)(C+\lambda')} = 0.$$

We know that

$$\frac{x^2}{A+\lambda} + \frac{y^2}{B+\lambda} + \frac{z^2}{C+\lambda} = 1 \quad \text{and} \quad \frac{x^2}{A+\lambda'} + \frac{y^2}{B+\lambda'} + \frac{z^2}{C+\lambda'} = 1.$$

Subtracting these equalities we obtain $(\lambda-\lambda')$ times the equality to prove. Since $\lambda \neq \lambda'$, we are ready. ∎

Regular surfaces swept out by a moving straight line are ruled surfaces. A bit more generally, we shall call a regular surface ruled, if every point of the surface has a neighborhood with a regular parameterization of the form

$$\mathbf{r}(u,v) = \gamma(u) + v\,\delta(u),$$

where γ is a smooth curve, called the directrix, δ is a nowhere zero vector field along γ. The straight lines $v \longmapsto \gamma(u_o) + v\,\delta(u_o)$ are the generators of the surface.

Theorem. The following propositions are equivalent for ruled surfaces:

(i) the normal vector field N is constant along the generators;

(ii) $\mathbf{r}_{uv}$ is tangential for the parameterization $\mathbf{r}(u) = \gamma(u) + v\,\delta(u)$;

(iii) the Gaussian curvature K is constant 0.

Proof. (i) $\Rightarrow$ (iii) If N is constant along the generators, then $L(\mathbf{r}_v) = -N_v = \mathbf{0} = 0\,\mathbf{r}_v$, thus generators are lines of curvature and the corresponding principal curvature is 0 everywhere. From this follows that the Gaussian curvature is 0.

(iii) $\Rightarrow$ (i) The normal section of a ruled surface in the direction of a generator is the generator itself. Hence, the normal curvature of the surface in the direction $\mathbf{r}_v$ of the generators is 0. If $\mathbf{r}_v$ were not a principal direction, then 0 would be strictly between the principal curvatures, in which case we would have $K < 0$. The contradiction shows that $\mathbf{r}_v$ is a principal direction at a given point, and thus $-N_v = L(\underline{r}_v) = 0\underline{r}_v = \underline{0}$, i.e. N is constant along the generators.

Remark. We have proved here that $K \leqslant 0$ for any ruled surface.

(iii) $\Longleftrightarrow$ (ii) According to the formula

$$K = \frac{\det \mathcal{B}}{\det \mathcal{G}},$$

$K = 0$ if and only if $\det \mathcal{B} = 0$. Since $\mathbf{r}_{vv} = 0$,

$$\det \mathcal{B} = \det \begin{pmatrix} \langle \mathbf{r}_{uu}, N\rangle & \langle \mathbf{r}_{uv}, N\rangle \\ \langle \mathbf{r}_{uv}, N\rangle & \langle \mathbf{r}_{vv}, N\rangle \end{pmatrix} = \det \begin{pmatrix} \langle \mathbf{r}_{uu}, N\rangle & \langle \mathbf{r}_{uv}, N\rangle \\ \langle \mathbf{r}_{uv}, N\rangle & 0 \end{pmatrix} = \langle \mathbf{r}_{uv}, N\rangle^2$$

and thus $\det \mathcal{B} = 0$ if and only if $\mathbf{r}_{uv}$ is orthogonal to N i.e. if $\mathbf{r}_{uv}$ is tangential. ∎

Definition. A ruled surface that satisfies one of the equivalent conditions of the previous theorem is called a developable surface.

Examples of *ruled but not developable surfaces:*

a) $\mathbf{r}(u,v) = (u, v, uv)$ (hyperbolic paraboloid);

b) $\mathbf{r}(u,v) = (a\cos u,\ b\sin u,\ 0) + v\,(-a\sin u,\ b\cos u,\ c)$

(one sheet hyperboloid)

c) $\mathbf{r}(u,v) = (0,0,u) + v\,(\cos u, \sin u,\ 0)$ (helicoid).

Examples of *developable surfaces:*

a) Cylinders over a curve. Let γ be a regular space curve, $\mathbf{v} \neq \mathbf{0}$ a vector nowhere tangent to γ. We define a cylinder over γ by the parameterization

$$\mathbf{r}(u,v) = \gamma(u) + v\,\mathbf{v}.$$

Since $\mathbf{r}_{uv} = \mathbf{0}$ is tangential, cylinders over a curve are developable.

b) Cones over a curve. Let γ be a regular curve, $\mathbf{p}$ be the position vector of a point not lying on any tangent to the curve. The cone over γ with vertex $\mathbf{p}$ is defined by the parameterization

$$\mathbf{r}(u,v) = v\,\gamma(u) + (1-v)\,\mathbf{p}\ .$$

The cone is regular only in the domain $v \neq 0$. The tangent plane of the cone is spanned by the vectors

$$\mathbf{r}_u(u,v) = v\,\gamma'(u) \quad \text{and} \quad \mathbf{r}_v(u,v) = \gamma(u) - \mathbf{p}.$$

Since

$$\mathbf{r}_{uv}(u,v) = \gamma'(u) = (1/v)\,\mathbf{r}_u(u,v),$$

cones are developable.

c) Tangential developables. Let γ be a curve of general type in $\mathbb{R}^3$. We show that the regular part of the surface swept out by the tangent lines of γ is developable. Indeed, the surface can be parameterized by

$$\mathbf{r}(u,v) = \gamma(u) + v\,\gamma'(u).$$

Partial derivatives of $\mathbf{r}$ are

$$\mathbf{r}_u(u,v) = \gamma'(u) + v\,\gamma''(u) \quad \text{and} \quad \mathbf{r}_v(u,v) = \gamma'(u).$$

Since γ is of general type, $\gamma'(u)$ and $\gamma''(u)$ are linearly independent, hence singularities of the surface of tangent lines are located along the generating curve γ.

Since $\mathbf{r}_{uv}(u,v) = \gamma''(u) = (1/v)\big(\mathbf{r}_u(u,v) - \mathbf{r}_v(u,v)\big)$ for $v \neq 0$, the regular part of the surface of tangent lines is developable. Surfaces of this type are called tangential developables.

As we proved in the theorem above, a ruled surface with Gaussian curvature equal to zero is developable. The following theorem states that in most cases the condition of being ruled follows from $K \equiv 0$.

Theorem. If the Gaussian curvature of a surface is 0 everywhere and the surface contains no planar point, then it is developable.

Proof. Gaussian curvature is positive at spherical points so the surface contains no umbilics. Therefore we may consider a parameterization $\mathbf{r}$ around any point $\mathbf{p}$ such that $\mathbf{p} = \mathbf{r}(0,0)$, coordinate lines are lines of curvature and the coordinate lines through p are unit speed curves (see unit 6). Suppose that $\mathbf{r}_u$ corresponds to the nonzero principal curvature $\kappa_1 = \kappa \neq 0$. Then

$$N_u = -\kappa\, \mathbf{r}_u \qquad \text{and} \qquad N_v = 0\, \mathbf{r}_v = \mathbf{0}\ .$$

The second equation shows that N is constant along v-coordinate lines. What we have to show is that v-coordinate lines are straight lines. For this purpose, it is enough to show that $\mathbf{r}$ is linear in v, i.e. $\mathbf{r}_{vv} = \mathbf{0}$. We prove this in a tricky way using the fact that the only vector which is perpendicular to each vectors of a basis is the zero vector. According to this proposition, it suffices to show that $\mathbf{r}_{vv}$ is orthogonal to the vectors N, $\mathbf{r}_u$, $\mathbf{r}_v$.

(i) $\mathbf{r}_{vv} \perp N$. This equation follows from

$$0 = \frac{\partial}{\partial v} \langle \mathbf{r}_v, N \rangle = \langle \mathbf{r}_{vv}, N \rangle + \langle \mathbf{r}_v, N_v \rangle = \langle \mathbf{r}_{vv}, N \rangle.$$

(ii) $\mathbf{r}_{vv} \perp \mathbf{r}_u$. Since lines of curvature are orthogonal,

$$0 = \frac{\partial}{\partial v} \langle \mathbf{r}_u, \mathbf{r}_v \rangle = \langle \mathbf{r}_{uv}, \mathbf{r}_v \rangle + \langle \mathbf{r}_u, \mathbf{r}_{vv} \rangle.$$

On the other hand,

$$\langle \mathbf{r}_{uv}, \mathbf{r}_v \rangle = \langle -\frac{\partial}{\partial v} \big((1/\kappa)\, N_u\big), \mathbf{r}_v \rangle = \frac{\partial}{\partial v} \big((1/\kappa)\big)\, \kappa \langle \mathbf{r}_u, \mathbf{r}_v \rangle + (1/\kappa)\langle N_{vu}, \mathbf{r}_v \rangle = 0.$$

Combining these two equalities we get $\langle \mathbf{r}_{uv}, \mathbf{r}_v \rangle = 0$.

(iii) $\mathbf{r}_{vv} \perp \mathbf{r}_v$. This will follow from the observation that v-coordinate lines are all parameterized by arc length, i.e. $\| \mathbf{r}_v \| \equiv 1$. We know that $\| \mathbf{r}_v(0,v) \| = 1$ by the construction of $\mathbf{r}$. We also have

$$\frac{\partial}{\partial u} \langle \mathbf{r}_v, \mathbf{r}_v \rangle = 2\langle \mathbf{r}_{uv}, \mathbf{r}_v \rangle = 0,$$

showing that $\langle \mathbf{r}_v, \mathbf{r}_v \rangle$ does not depend on u. Thus,

$$\| \mathbf{r}_v(u,v) \| = \| \mathbf{r}_v(0,v) \| = 1 \text{ for every } u, v.$$

Now differentiating with respect to v,

$$0 = \frac{\partial}{\partial v} \langle \mathbf{r}_v, \mathbf{r}_v \rangle = 2\langle \mathbf{r}_{vv}, \mathbf{r}_v \rangle.$$

This completes the proof. ∎

We finish the investigation of developable surfaces with a structure theorem stating that every developable surface is made up of pieces of cylinders, cones and tangential developables.

Theorem. Let $\mathbf{r}\colon [a,b]\times[c,d] \longrightarrow \mathbb{R}^3$ be a developable surface without planar points and suppose that the parameterization $\mathbf{r}$ of the surface is the one we used in the proof of the previous theorem. Then there exists a nowhere dense closed subset A of [a,b] the complement of which is a union of open intervals $[a,b]\setminus A = I_1 \cup I_2 \cup \ldots$ such that the restriction of $\mathbf{r}$ onto $I_n \times [c,d]$ is a part of a cylinder or cone or a tangential developable.

Proof. As it was proved above, $\mathbf{r}$ has the form $\mathbf{r}(u,v) = \mathbf{a}(u) + v\mathbf{b}(u)$, where $\mathbf{b}(u)$ is a unit vector field along the curve $\mathbf{a}(u)$. We have

$$\mathbf{r}_u(u,v) = \mathbf{a}'(u) + v\, \mathbf{b}'(u) \qquad \mathbf{r}_v(u,v) = \mathbf{b}(u) \qquad \mathbf{r}_{uv}(u,v) = \mathbf{b}'(u).$$

By the definition of developable surfaces, $\mathbf{r}_{uv}(u,v) = \mathbf{b}'(u)$ must be tangential to the surface, i.e. it lies in the plane spanned by $\mathbf{r}_u$ and $\mathbf{r}_v$. $\mathbf{r}_u$ is orthogonal to $\mathbf{r}_v$ as they are lines of curvature, furthermore, $\mathbf{r}_{uv}$ is also orthogonal to $\mathbf{r}_v$ since $0 = \langle\mathbf{b}(u),\mathbf{b}(u)\rangle' = 2\langle\mathbf{b}(u),\mathbf{b}'(u)\rangle$. For there is only one direction in a plane which is orthogonal to a given nonzero vector, $\mathbf{r}_u$ and $\mathbf{r}_{uv}$ must be parallel: $\mathbf{b}'(u)\|\mathbf{a}'(u)+v\mathbf{b}'(u)$, or equivalently, $\mathbf{b}'(u)\|\mathbf{a}'(u)$. Hence, $\mathbf{b}'(u) = c(u)\,\mathbf{a}'(u)$ for some function $c:[a,b]\longrightarrow\mathbb{R}$. Now let A be the set of those roots of $c\,c'$ in $[a,b]$, which do not have a neighborhood consisting of only roots of $c\,c'$. A is closed and nowhere dense in $[a,b]$. If $[a,b]\setminus A$ is the union of the disjoint open intervals $I_1, I_2, \ldots$, then for the restriction of c onto I_n, we have one of the following possibilities:

(i) the restriction is identically 0;

(ii) the restriction is a nonzero constant;

(iii) the restriction is strictly monotone and nowhere zero.

In the first case, $\mathbf{b}'(u) = 0$ and thus $\mathbf{b}$ is constant on I_n, thus the restriction of $\mathbf{r}$ onto $I_n\times[c,d]$ is a part of a cylinder.

In the second case, the point $\mathbf{p}(u) = \mathbf{a}(u)-(1/c)\mathbf{b}(u)$ does not depend on u. Indeed, $\mathbf{p}'(u) = \mathbf{a}'(u)-(1/c)\mathbf{b}'(u) = \mathbf{0}$. Furthermore, the point $\mathbf{p}$ lies on every generator of the surface, so this case serves a part of a cone.

Finally, consider the curve $\gamma(u) = \mathbf{a}(u)-(1/c(u))\mathbf{b}(u)$ for the last case. As

$$\gamma'(u) = \mathbf{a}'(u) - (1/c(u))'\mathbf{b}(u) - (1/c(u))\mathbf{b}'(u) = -(1/c(u))'\mathbf{b}(u)\|\mathbf{b}(u),$$

the tangent of γ at $\gamma(u)$ coincides with the generator $v\longmapsto\mathbf{r}(u,v)$ of the surface. γ is of general type, since $\gamma'(u)\times\gamma''(u)=\big((1/c(u))'\big)^2\mathbf{b}(u)\times\mathbf{b}'(u) = \big((1/c(u))'\big)^2 c(u)\,\mathbf{b}(u)\times\mathbf{a}'(u)\neq\mathbf{0}$. We conclude that in the third case the restriction of $\mathbf{r}$ onto $I_n\times[c,d]$ is a part of the tangential developable generated by the curve of general type γ. ∎

UNIT 7. THE FUNDAMENTAL EQUATIONS OF HYPERSURFACE THEORY

Gauss frame of a parameterized hypersurface, formulae for the partial derivatives of the Gauss frame vector fields, Christoffel symbols, Gauss and Codazzi-Mainardi equations, fundamental theorem of hypersurfaces, "Theorema Egregium", components of the curvature tensor, tensors in linear algebra, tensor fields over a hypersurface, curvature tensor.

Now we derive some formulae for hypersurfaces. Consider a regular parameterized hypersurface $\mathbf{r}:\Omega\longrightarrow \mathbb{R}^{n+1}$. The partial derivatives $\mathbf{r}_1,\dots,\mathbf{r}_n$ define a basis of the tangent space of the hypersurface at each point. If we add to these vectors the normal vector of the hypersurface, we get a basis of $\mathbb{R}^{n+1}$ at each point of the hypersurface. The system of the vector fields $\mathbf{r}_1,\dots,\mathbf{r}_n,N$ along $\mathbf{r}$ is called the Gauss frame of the hypersurface. Gauss frame plays similar role in the theory of hypersurfaces as Frenet frame does in curve theory. Similarity is not complete however, since a Gauss frame is much more dependent on the parameterization. Nevertheless, in the same way as for Frenet frames, it is important to know how the derivatives of the frame vector fields with respect to the parameters can be expressed as a linear combination of the frame vectors. For this we have to determine the coefficients Γ_{ij}^k, α_{ij}, β_j^k, γ_j in the expressions

$$\mathbf{r}_{ij} = \sum_k \Gamma_{ij}^k \mathbf{r}_k + \alpha_{ij}N, \qquad N_j = \sum_k \beta_j^k \mathbf{r}_k + \gamma_j N. \qquad (*)$$

Let us begin with the simple observation that since N_j is known to be tangential, and $N_j = -L(\mathbf{r}_j)$, where L is the Weingarten map,

$$\boxed{\gamma_j = 0 \text{ for all } j}$$

and $(-\beta_j^k)_{j,k=1}^n$ is the matrix $\mathcal{L} = \mathcal{B}\,\mathcal{G}^{-1}$ of the Weingarten map with respect to the basis $\mathbf{r}_1,\dots,\mathbf{r}_n$. Denote by g_{ij} and b_{ij} the entries of the first and second fundamental forms as usual, and denote by g^{ij} the components of the inverse matrix of the matrix of the first fundamental form. (Attention! Entries of $\mathcal{G}$ and $\mathcal{G}^{-1}$ are distinguished by the position of indices.) Then

$$\boxed{\beta_j^k = -\sum_i b_{ji}g^{ik}}.$$

Taking the dot product of the first equation of (*) with N we gain the equality $\langle \mathbf{r}_{ij}, N\rangle = \alpha_{ij}$ and since $\langle \mathbf{r}_{ij}, N\rangle = b_{ij}$,

$$\boxed{\alpha_{ij} = b_{ij} \quad \text{for all } i,j}\,.$$

There is only one question left: what are the coefficients Γ_{ij}^{k} equal to? Let us take the dot product of the first equation of (*) with $\mathbf{r}_\ell$

$$\langle \mathbf{r}_{ij}, \mathbf{r}_\ell\rangle = \sum_k \Gamma_{ij}^{k} \langle \mathbf{r}_k, \mathbf{r}_\ell\rangle = \sum_k \Gamma_{ij}^{k} g_{k\ell} \quad ,$$

or denoting the dot product $\langle \mathbf{r}_{ij}, \mathbf{r}_\ell\rangle$ shortly by $\Gamma_{ij\ell}$,

$$\Gamma_{ij\ell} = \sum_k \Gamma_{ij}^{k} g_{k\ell} \quad .$$

The coefficients Γ_{ij}^{k} and Γ_{ijk} are called the <u>Christoffel symbols of first and second type</u> respectively. The last equation shows how to express Christoffel symbols of second type with the help of Christoffel symbols of first type. It can also be used to express Christoffel symbols of first type in terms of secondary Christoffel symbols. Indeed, multiplying the equation with $g^{\ell s}$, taking sum for ℓ and using $\sum_\ell g_{k\ell} g^{\ell s} = \delta_k^\ell$ (← Kronecker delta), we get

$$\sum_\ell \Gamma_{ij\ell} g^{\ell s} = \sum_\ell \sum_k \Gamma_{ij}^{k} g_{k\ell} g^{\ell s} = \sum_k \Gamma_{ij}^{k} \delta_k^s = \Gamma_{ij}^{s} \,.$$

Now let us try to determine Christoffel symbols of second type. Differentiating the equality $g_{ij} = \langle \mathbf{r}_i, \mathbf{r}_j\rangle$ with respect to the k-th variable and then permuting the role of indices i,j,k we get the equalities

$$\begin{aligned} g_{ij,k} &= \langle \mathbf{r}_{ik}, \mathbf{r}_j\rangle + \langle \mathbf{r}_i, \mathbf{r}_{jk}\rangle \\ g_{jk,i} &= \langle \mathbf{r}_{ji}, \mathbf{r}_k\rangle + \langle \mathbf{r}_j, \mathbf{r}_{ki}\rangle \\ g_{ki,j} &= \langle \mathbf{r}_{kj}, \mathbf{r}_i\rangle + \langle \mathbf{r}_k, \mathbf{r}_{ij}\rangle \,. \end{aligned}$$

Solving this linear system of equations for the secondary Christoffel symbols standing on the right hand side, we obtain

$$\Gamma_{ijk} = \langle \mathbf{r}_{ij}, \mathbf{r}_k\rangle = \frac{1}{2}\left(g_{ik,j} + g_{jk,i} - g_{ij,k}\right)$$

and

$$\boxed{\Gamma_{ij}^{k} = \sum_\ell \Gamma_{ij\ell} g^{\ell k} = \sum_\ell \frac{1}{2}\left(g_{i\ell,j} + g_{j\ell,i} - g_{ij,\ell}\right) g^{\ell k}} \quad .$$

Observe that the Christoffel symbols *depend only on the first fundamental form* of the hypersurface.

Now we ask the following question. Suppose we are given $2n^2$ smooth functions g_{ij}, b_{ij} i,j=1,2,...,n on an open domain Ω of $\mathbb{R}^{n+1}$. When can we find a parameterized hypersurface $\mathbf{r}:\Omega\longrightarrow \mathbb{R}^{n+1}$ with fundamental forms $\mathcal{G} = (g_{ij})$ and $\mathcal{B} = (b_{ij})$. We have some obvious restrictions on the functions g_{ij} and b_{ij}. First, $g_{ij} = g_{ji}$, $b_{ij} = b_{ji}$, and since $\mathcal{G}$ is the matrix of a positive definite bilinear form, the determinants of the corner submatrices $(g_{ij})_{i,j=1}^{k}$

must be positive for k = 1,...,n. However, the examples we have show that these conditions are not enough to guarantee the existence of a hypersurface. For example, if $\mathcal{G}$ is the identity matrix everywhere, while $\mathcal{B} = f\,\mathcal{G}$ for some function on Ω, then the hypersurface (if exists) consists of umbilics. We know however that if a surface consists of umbilics, then the principal curvatures are constant, so although our choice of $\mathcal{B}$ and $\mathcal{G}$ satisfies all the conditions we have listed so far, it does not correspond to a hypersurface unless f is constant. So there must be some further relations between the components of $\mathcal{B}$ and $\mathcal{G}$. Our plan to find some of these correlations is the following. Let us express $\mathbf{r}_{ijk}$ and $\mathbf{r}_{ikj}$ as a linear combination of the Gauss frame vectors. The coefficients we get are functions of the entries of the first and second fundamental forms. For $\mathbf{r}_{ijk} = \mathbf{r}_{ikj}$, the corresponding coefficients in the expressions for these vectors must be equal and it can be hoped that this way we arrive at further non-trivial relations between $\mathcal{G}$ and $\mathcal{B}$. This was the philosophy, and now let us get down to work.

$$\mathbf{r}_{ijk} = \Big(\sum_{\ell} \Gamma_{ij}^{\ell} \mathbf{r}_{\ell} + b_{ij}N \Big)_{,k} = \sum_{\ell} \Big(\Gamma_{ij,k}^{\ell}\mathbf{r}_{\ell} + \Gamma_{ij}^{\ell} \mathbf{r}_{\ell k} \Big) + b_{ij,k}N + b_{ij}N_{k} =$$

$$= \sum_{\ell} \Big(\Gamma_{ij,k}^{\ell}\mathbf{r}_{\ell} + \Gamma_{ij}^{\ell} \Big(\sum_{s} \Gamma_{\ell k}^{s} \mathbf{r}_{s} + b_{\ell k}N\Big)\Big) + b_{ij,k}N - b_{ij} \sum_{\ell s}\sum b_{ks}g^{s\ell}\mathbf{r}_{\ell} =$$

$$= \sum_{\ell} \Big(\Gamma_{ij,k}^{\ell} + \sum_{s} \Gamma_{ij}^{s} \Gamma_{sk}^{\ell} - b_{ij} \sum_{s} b_{ks}g^{s\ell}\Big) \mathbf{r}_{\ell} + \Big(b_{ij,k} + \sum_{\ell} \Gamma_{ij}^{\ell} b_{\ell k}\Big)N .$$

Comparing the coefficient of $\mathbf{r}_{\ell}$ in $\mathbf{r}_{ijk}$ and $\mathbf{r}_{ikj}$, we obtain

$$\boxed{\Gamma_{ij,k}^{\ell} - \Gamma_{ik,j}^{\ell} + \sum_{s} \Big(\Gamma_{ij}^{s} \Gamma_{sk}^{\ell} - \Gamma_{ik}^{s} \Gamma_{sj}^{\ell}\Big) = \sum_{s} \Big(b_{ij} b_{ks} - b_{ik} b_{js}\Big) g^{s\ell}}$$

while comparison of the coefficient of N gives

$$\boxed{b_{ij,k} - b_{ik,j} = \sum_{\ell} \Gamma_{ik}^{\ell} b_{\ell j} - \sum_{\ell} \Gamma_{ij}^{\ell} b_{\ell k}}\,.$$

The first n^4 equations (we have an equation for all i,j,k,ℓ), are the Gauss equations for the hypersurface. The second family of n^3 equations are the Codazzi-Mainardi equations.

Exercise. Express the second order derivatives N_{ij} and N_{ji} as a linear combination of the Gauss frame vectors. Compare the corresponding coefficients and prove that their equality follows from the Gauss and Codazzi-Mainardi equations.

The exercise points out that a similar try to derive new relations between $\mathcal{G}$ and $\mathcal{B}$ does not lead to really new results. This is no wonder, since the Gauss and Codazzi-Mainardi equations together with the previously listed obvious conditions on $\mathcal{G}$ and $\mathcal{B}$ form a complete system of necessary and sufficient conditions for the existence of a hypersurface with fundamental forms $\mathcal{G}$ and $\mathcal{B}$.

Theorem. (Fundamental theorem of hypersurfaces). Let $\Omega \subset \mathbb{R}^n$ be an open connected and simply connected subset of $\mathbb{R}^n$ (e.g. an open ball or cube), and suppose that we are given two smooth n by n matrix valued functions $\mathcal{G}$ and $\mathcal{B}$ on Ω such that $\mathcal{G} = (g_{ij})$ and $\mathcal{B} = (b_{ij})$ assign to every point a symmetric matrix, $\mathcal{G}$ gives the matrix of a positive definite bilinear form. In this case, if the functions Γ_{ij}^{k} derived from the components of $\mathcal{G}$ according to the above formulae satisfy the Gauss and Codazzi-Mainardi equations, then there exists a regular parameterized hypersurface $\mathbf{r}:\Omega\longrightarrow \mathbb{R}^{n+1}$ for which the matrix representations of the first and second fundamental forms are $\mathcal{G}$ and $\mathcal{B}$ respectively. Furthermore, this hypersurface is unique up to rigid motions of the whole space. Namely, if $\mathbf{r}_1$ and $\mathbf{r}_2$ are two such hypersurfaces, then there exists an isometry (=distance preserving bijection) $\Phi:\mathbb{R}^{n+1}\longrightarrow \mathbb{R}^{n+1}$ for which $\mathbf{r}_2 = \Phi\circ\mathbf{r}_1$.

Let us denote the expressions standing on the left hand sides of the Gauss equations by

$$\boxed{R_{ijk}^{\ell} := \Gamma_{ij,k}^{\ell} - \Gamma_{ik,j}^{\ell} + \sum_s \left(\Gamma_{ij}^{s}\,\Gamma_{sk}^{\ell} - \Gamma_{ik}^{s}\,\Gamma_{sj}^{\ell}\right)}\,.$$

Then Gauss equations can be abbreviated writing

$$R_{ijk}^{\ell} = \sum_s \left(b_{ij}\,b_{ks} - b_{ik}\,b_{js}\right) g^{s\ell}.$$

Let us multiply this equation by $g_{\ell m}$ and take a sum for ℓ

$$\sum_\ell R_{ijk}^{\ell}\,g_{\ell m} = \sum_\ell\sum_s \left(b_{ij}\,b_{ks} - b_{ik}\,b_{js}\right) g^{s\ell}\,g_{\ell m} =$$
$$= \sum_s \left(b_{ij}\,b_{ks} - b_{ik}\,b_{js}\right) \delta_m^s = \left(b_{ij}\,b_{km} - b_{ik}\,b_{jm}\right).$$

Introducing the functions $\boxed{R_{imjk} := \sum_\ell R_{ijk}^{\ell}\,g_{\ell m}}$, we may write

$$\boxed{R_{imjk} = \left(b_{ij}\,b_{km} - b_{ik}\,b_{jm}\right)}\,.$$

Let us observe, that the functions R_{imjk} can be expressed in terms of the first fundamental form $\mathcal{G}$.

Corollary. (Theorema Egregium) The Gaussian curvature of a regular parameterized surface in $\mathbb{R}^3$ can be expressed in terms of the first fundamental form as follows

$$K = \frac{R_{1212}}{\det \mathcal{G}}\,.$$

Theorema Egregium is one of those theorems of Gauss he was very proud of.

The surprising fact is not the actual form of this formula but the mere existence of a formula that expresses the Gaussian curvature in terms of the first fundamental form. The geometrical meaning of the existence of such a formula is that the *Gaussian curvature does not change when we bend the surface* (although principal curvatures do change in general!).

Definition. Let $\mathbf{r}:\Omega\longrightarrow \mathbb{R}^{n+1}$ be a hypersurface. Consider the mapping R that assigns four tangential vector fields $\mathbf{X} = \sum_i X^i\mathbf{r}_i$, $\mathbf{Y} = \sum_i Y^i\mathbf{r}_i$, $\mathbf{Z} = \sum_i Z^i\mathbf{r}_i$, $\mathbf{W} = \sum_i W^i\mathbf{r}_i$ a function according to the formula

$$R(X,Y;Z,W) = \sum_i \sum_m \sum_j \sum_k R_{imjk}X^iY^mZ^jW^k.$$

We shall call R the curvature tensor of the hypersurface, the functions R_{imjk} the components of the curvature tensor.

Let us briefly recall some definition from linear algebra, concerning tensors.

Let V be a vector space (over $\mathbb{R}$). The set V^* of linear functions form a vector space with respect to the operations

$$(\ell_1+\ell_2)(\mathbf{v}) := \ell_1(\mathbf{v}) + \ell_2(\mathbf{v}), \qquad (\lambda\,\ell)(\mathbf{v}) := \lambda\big(\ell(\mathbf{v})\big).$$

The vector space V^* of linear functions on V is called the dual space of V. If V is finite dimensional and $\mathbf{e}_1,\dots,\mathbf{e}_n$ is a basis of V, then we may consider the linear functions $\mathbf{e}^1,\dots,\mathbf{e}^n \in V^*$ defined by $\mathbf{e}^i(\mathbf{e}_j) = \delta^i_j$. It is not difficult to prove that these linear functions form a basis of V^* called the dual basis of the basis $\mathbf{e}_1,\dots,\mathbf{e}_n$. As a consequence we get that dim V = dim V^* for finite dimensional vector spaces. A tensor of valency (/order /type) (k,ℓ) over V is a multilinear function

$$T : V^*\times\dots\times V^*\times V\times\dots\times V \longrightarrow \mathbb{R}$$

defined on the Cartesian product of k copies of V^* and ℓ copies of V. "Multilinear" means that fixing all but one variables, we obtain a linear function of the free variable. Denote by $T^{(k,\ell)}V$ the set of tensors of valency (k,ℓ). The sum of two tensors of order (k,ℓ) and the scalar multiple of a tensor are tensors of the same order, hence the set of tensors of a given valency form a vector space. If $\mathbf{e}_1,\dots,\mathbf{e}_n$ is a basis of V, then every tensor T is uniquely determined by its values on basis vector combinations, i.e. by the numbers

$$T^{i_1\dots i_k}_{j_1\dots j_\ell} = T(\mathbf{e}^{i_1},\dots,\mathbf{e}^{i_k};\mathbf{e}_{j_1},\dots,\mathbf{e}_{j_\ell}),$$

which are called the components of the tensor T with respect to the basis

$\mathbf{e}_1,\dots,\mathbf{e}_n$. Since any $(\dim V)^{(k+\ell)}$ numbers $T^{i_1\dots i_k}_{j_1\dots j_\ell}$ correspond to a tensor, $\dim T^{(k,\ell)}V = (\dim V)^{(k+\ell)}$.

Now let us consider a regular parameterized hypersurface M, $\underline{r}:\Omega \longrightarrow \mathbb{R}^{n+1}$. A <u>tensor field of valency</u> (k,ℓ) over M is a mapping T that assigns to every point $\underline{u} \in \Omega$ a tensor of valency (k,ℓ) over the tangent space of M at $\underline{r}(\underline{u})$. $T(\underline{u})$ is uniquely determined by its components $T^{i_1\dots i_k}_{j_1\dots j_\ell}(\underline{u})$ with respect to the basis $\mathbf{r}_1(\underline{u}),\dots,\mathbf{r}_n(\underline{u})$. The functions $\underline{u} \longmapsto T^{i_1\dots i_k}_{j_1\dots j_\ell}(\underline{u})$ are called the <u>components of the tensor field</u> T. T is said to be a <u>smooth tensor field</u> if its components are smooth.

Examples.

- Function on M are tensor fields of valency (0,0).
- Tangential vector fields are tensor fields of valency (1,0) (V is isomorphic to V^{**} in a natural way).
- The first and second fundamental forms of a hypersurface are tensor fields of valency (0,2).
- The mapping that assigns to every point of a hypersurface the Weingarten map at that point is a tensor of valency (1,1). (The linear space of $V \longrightarrow V$ linear mappings is isomorphic to $T^{(1,1)}V$ in a natural way.)
- Let f be a smooth function on M. Consider the tensor field of valency (0,1) defined on a tangent vector X to be the derivative of f in the direction X. This tensor field is the differential of f.
- The curvature tensor is a tensor field of valency (0,4).

The curvature tensor is one of the most basic objects of study in differential geometry. In the previous computations the curvature tensor came across like a rabbit from a cylinder. To understand its real meaning, we shall introduce the curvature tensor in a more natural way in a more general framework, in the framework of Riemannian manifolds. For this purpose, we have to get acquainted with some fundamental definitions and constructions. This will be the goal of the following units.

Unit 8. Topological and Differentiable Manifolds

The configuration space of a mechanical system, examples; the definition of topological and differentiable manifolds, smooth maps and diffeomorphisms; Lie groups, embedded submanifolds in $\mathbb{R}^n$, examples, Whitney's theorem, classification of closed 2-manifolds.

As the motion of a particle in $\mathbb{R}^3$ corresponds to a parameterized space curve, a motion of a system of n points can be described by n parameterized curves $\mathbf{x}_i:[a,b]\longrightarrow \mathbb{R}^3$ $i = 1,2,\ldots,n$.

Putting these mappings together, we obtain a curve

$$(\mathbf{x}_1,\ldots,\mathbf{x}_n):[a,b]\longrightarrow \mathbb{R}^3\times\mathbb{R}^3\times\ldots\times\mathbb{R}^3 \text{ (n times)}$$

in the direct product of n copies of $\mathbb{R}^3$, the projections of which on the i-th factor of the product is just the curve $\mathbf{x}_i$.

For there is a one-to-one correspondence between points of the product $\mathbb{R}^3\times\mathbb{R}^3\times\ldots\times\mathbb{R}^3$ (n times) and the possible configurations of n points in the space, we shall call the direct product of n copies of $\mathbb{R}^3$ the *configuration space* of the system of n points.

In general, the configuration space of a mechanical system is the set of all of its possible positions, equipped with some natural additional structures such as topology or the structure of a differentiable manifold (see later).

The advantage of introducing the configuration space is that the motion of the system can be interpreted as one single curve in the configuration space instead of a set of space curves.

Non-trivial examples can be obtained by putting some constraints on a system of n points. For example, some pairs of points can be connected by a rigid rod, some points can be fixed or forced to move along a line or a surface. Further constraints can be obtained by specifying the type of joint at the points where two or more rods meet.

The configuration space of a system of n points with constraints is a subspace of $\mathbb{R}^{3n}$ and it is quite natural to furnish it with the subspace topology inherited from $\mathbb{R}^{3n}$.

Examples.

i) The configuration space of the planar pendulum is the circle S^1.

ii) The configuration space of the spherical pendulum is the two-dimensional sphere S^2.

iii) The configuration space of a planar double pendulum is the direct product of two circles, i.e. the torus $T^2 = S^1 \times S^1$.

iv) The configuration space of a spherical double pendulum is the direct product of two spheres $S^2 \times S^2$.

v) A rigid segment in the plane has for its configuration space the direct product $\mathbb{R}^2 \times S^1$, which is homeomorphic to the open solid torus.

As we see in the above examples, the configuration space of a mechanical system is not necessarily homeomorphic to a linear space, but in each case the points of the configuration space have a neighborhood homeomorphic to an open ball.

In the following chain of definitions we *fix a positive integer* n.

Definition. Let X be an arbitrary set. A local parameterization of X is an injective mapping $\varphi : \Omega \longrightarrow X$ from an open subset Ω of $\mathbb{R}^n$ onto a subset of X.

The inverse $\varphi^{-1}: \varphi(\Omega) \longrightarrow \Omega$ of such a parameterization is called a chart because through φ^{-1} the region im $\varphi \subset X$ is "charted" on $U \subset \mathbb{R}^n$, just as a region of the earth is charted on a topographic or a political map. φ^{-1} is also called a coordinate system because through φ^{-1} each point $p \in$ im φ corresponds to an n-tuple of real numbers, the coordinates of p.

An atlas on X is a collection of charts $\mathcal{A} = \{ \varphi_i : i \in I \}$ such that every point is represented in at least one chart i.e. $\bigcup_{i \in I} \text{dom}\ \varphi_i = X$.

Two charts $\varphi : X \hookrightarrow U$ and $\psi : X \hookrightarrow V$ are said to be $\mathcal{C}^r$-compatible if the domains $\varphi(\text{dom}\ \psi \cap \text{dom}\ \varphi)$ and $\psi(\text{dom}\ \psi \cap \text{dom}\ \varphi)$ of the "transit" mappings $\psi\circ\varphi^{-1}$ and $\varphi\circ\psi^{-1}$ are open subsets of $\mathbb{R}^n$, and $\psi\circ\varphi^{-1}$ and $\varphi\circ\psi^{-1}$ are r times continuously differentiable. (A mapping is 0 times continuously differentiable if it is continuous).

An atlas is $\mathcal{C}^r$-compatible if any two charts in the atlas are $\mathcal{C}^r$-compatible.

An atlas $\mathcal{A}$ on a set X defines a topology on X as follows

Let $U \subset X$ be open if and only if $\varphi\ (U \cap \text{dom}\ \varphi)$ is open in $\mathbb{R}^n$ with respect to any chart φ from $\mathcal{A}$.

Proposition. The family of open sets yields a topology on X.

Proof. The statement follows directly from the following set theoretical identities.

i) $\varphi\ (\emptyset \cap \text{dom}\ \varphi) = \emptyset$

ii) $\varphi\ (X \cap \text{dom}\ \varphi) = \text{im}\ \varphi$

iii) $\varphi\ (U \cap V \cap \text{dom}\ \varphi) = \varphi\ (U \cap \text{dom}\ \varphi) \cap \varphi\ (V \cap \text{dom}\ \varphi)$

iv) $\varphi\left(\left(\bigcup_{i\in I} U_i\right)\cap \operatorname{dom}\varphi\right) = \bigcup_{i\in I} \varphi(U_i \cap \operatorname{dom}\varphi)$. ▮

Definition. An n-dimensional topological manifold is a pair $(X,\mathcal{A})$ consisting of a point set X and a $\mathcal{C}^0$-compatible atlas $\mathcal{A}$ on it, such that the topology induced by the atlas on X satisfies the following two conditions

i) for any two distinct points $x,\ y \in X$, one can find two disjoint neighborhoods of x and y (i.e. X is a Hausdorff space);

ii) there exists a countable family of charts $\varphi_1,\varphi_2,\varphi_3,\ldots \in \mathcal{A}$, the domain of which cover X (X is a second countable topological space).

Remark. In physics, the dimension of the configuration space of a mechanical system (provided that it is a manifold) is called the number of degrees of freedom.

We say that a topological manifold is a $\mathcal{C}^r$-manifold if the atlas $\mathcal{A}$ of it is $\mathcal{C}^r$-compatible. Two atlases are *equivalent* or define the same $\mathcal{C}^r$-manifold structure on X if their union also consists of $\mathcal{C}^r$-compatible charts. It is clear that each equivalence class of atlases contains a unique *maximal atlas*.

We shall mainly be interested in $\mathcal{C}^\infty$-manifolds which will also be called smooth or differentiable manifolds.

Examples.

(i) $\mathbb{R}^n$ equipped with the atlas consisting of only one chart, the identity mapping of $\mathbb{R}^n$, is an n-dimensional differentiable manifold.

(ii) Open subsets $U \subset X$ of an n-dimensional manifold $(X,\mathcal{A})$ become n-dimensional manifolds with the atlas $\{\varphi|_{\operatorname{dom}(\varphi)\cap U} : \varphi \in \mathcal{A}\}$.

(iii) If $(X_1,\mathcal{A}_1)$ and $(X_2,\mathcal{A}_2)$ are two manifolds of dimension n and m respectively, then the product space $X_1 \times X_2$ has a natural (n+m)-dimensional manifold structure given by the atlas

$$\{(\varphi_1,\varphi_2):\operatorname{dom}(\varphi_1) \times \operatorname{dom}(\varphi_2)\longrightarrow \mathbb{R}^{n+m} : \varphi_1\in\mathcal{A}_1,\ \varphi_2\in\mathcal{A}_2\}.$$

(iv) We have introduced the topology on the Grassmann manifolds Gr(n,k) in Unit 2. The topology of these spaces comes from a k(n-k)-dimensional differentiable manifold structure. We construct a chart $\varphi_{\mathcal{B}}$ on Gr(n,k) to every ordered basis $\mathcal{B} = \{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ of $\mathbb{R}^n$. Let us denote by V the subspace spanned by the first k vectors of $\mathcal{B}$ and by W the subspace spanned by the last (n-k) vectors. It is clear that $\mathbb{R}^n = V \oplus W$. Denote by $\pi:\mathbb{R}^n\longrightarrow V$ the projection of $\mathbb{R}^n$ onto V along W. The chart $\varphi_{\mathcal{B}}$ will be defined on the set

$$\operatorname{dom}(\varphi_{\mathcal{B}}) = \{L \in Gr(n,k) : L \cap W = \{\mathbf{0}\}\}.$$

$\varphi_{\mathcal{B}}$ assigns to $L \in \operatorname{dom}(\varphi_{\mathcal{B}})$ a $k\times(n-k)$ matrix in the following way. The restriction of π onto L is an isomorphism between L and V. The preimages of the vectors $\mathbf{x}_1,\ldots,\mathbf{x}_k$ yield a basis of L which has the form

$$(\pi|_L)^{-1}(\mathbf{x}_i) = \mathbf{x}_i + \sum_{j=k+1}^{n} \alpha_{ij}\, \mathbf{x}_j.$$

It is clear that setting $\varphi_{\mathcal{B}}(L)$ to be equal to the matrix of coefficients (α_{ij}), $i=1,\dots,k$; $j=k+1,\dots,n$, we obtain a bijection between $\mathrm{dom}(\varphi_{\mathcal{B}})$ and the set of all $k\times(n-k)$ matrices. The family of all charts of the form $\varphi_{\mathcal{B}}$ is a $\mathcal{C}^{\infty}$-compatible atlas on $Gr(n,k)$.

$Gr(n,k)$ is a compact manifold, it can be covered by a finite number of charts. Indeed we get a finite atlas on $Gr(n,k)$ if we let $\mathcal{B}$ run through different permutations of the standard basis of $\mathbb{R}^n$.

The Grassmann manifold $Gr(n+1,1)$ is the n-dimensional projective space. The geometrical way to introduce projective spaces is the following. We take an n-dimensional Euclidean space and join to it a collection of extra points, called ideal points or points at infinity, in such a way, that we attach one point at infinity to each straight line and two straight line gets the same point at infinity if and only if they are parallel. If we put the n-dimensional space into the (n+1)-dimensional one and fix a point O outside it, then every straight line through O intersects the projective closure of the n-dimensional Euclidean space in a unique ordinary or ideal point and this is the natural correspondence between the two ways of introducing projective spaces.

A mapping $f: X \longrightarrow Y$ from a differentiable manifold $(X,\mathcal{A})$ into the differentiable manifold $(Y,\mathcal{B})$ is said to be smooth if for any two charts $\varphi\in\mathcal{A}$ and $\psi \in \mathcal{B}$, the mapping $\psi\circ f\circ\varphi^{-1}$ is smooth. The map f is a diffeomorphism if it is a bijection and both f and f^{-1} are smooth.

Two differentiable manifolds are diffeomorphic if there is a diffeomorphism between them.

Definition. A Lie group is a differentiable manifold G with a group operation such that the mapping

$$G \times G \longrightarrow G, \quad (x,y) \longmapsto x\, y^{-1}$$

is differentiable.

Example. $Gl(n,\mathbb{R})$ and $Gl(n,\mathbb{C})$ are open subset in the linear spaces of all $n\times n$ real/complex matrices, hence they have a differentiable manifold structure. They also have a group structure, which is smooth since the entries of the quotient of two matrices are rational functions of the entries of the original matrices and rational functions are smooth. This way, general linear groups are Lie groups.

The following theorem explains why the configuration space of a system of n points with constraints so often happens to be a manifold.

Theorem. Let $F:\mathbb{R}^n \longrightarrow \mathbb{R}^k$ be a smooth mapping, the image of which contains $\mathbf{0} \in \mathbb{R}^k$. Consider the preimage of the point $\mathbf{0}$

$$X = \{ \mathbf{x} \in \mathbb{R}^n : F(\mathbf{x}) = \mathbf{0} \}.$$

Let us suppose that the gradient vectors

$$\operatorname{grad} f_i(\mathbf{x}) = \left(\frac{\partial f_i}{\partial x_1}(\mathbf{x}), \frac{\partial f_i}{\partial x_2}(\mathbf{x}), \dots, \frac{\partial f_i}{\partial x_n}(\mathbf{x}) \right)$$

of the coordinate functions of $F = (f_1, f_2, \dots f_n)$ are linearly independent at each point $\mathbf{x}$ of X.

Then X is an (n-k)-dimensional topological manifold, furthermore, there is a well-defined differentiable manifold structure on X.

Remark. The condition on the independence of the gradient vectors of the coordinate functions is essential. By a theorem due to Whitney, for any closed set $C \subset \mathbb{R}^n$ there exists a smooth function f on $\mathbb{R}^n$ such that $C = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x})=0\}$.

Proof. Let us recall a fundamental result of multivariable calculus.

The Inverse Function Theorem. If $\mathbf{x}$ is a point in the domain of a smooth function $\tilde{F} = (f_1, f_2, \dots, f_n): U \longrightarrow \mathbb{R}^n$, defined on an open subset of $\mathbb{R}^n$, and the gradient vectors of the coordinate functions $f_1, f_2, \dots, f_n$ of $\tilde{F}$ are linearly independent at the point $\mathbf{x}$, then there exists an open neighborhood $V \subset U$ of $\mathbf{x}$ such that $\tilde{F}|_V$ is a diffeomorphism between V and $\tilde{F}(V)$. In addition, $\tilde{F}(V)$ is an open subset of $\mathbb{R}^n$.

The linear space $\mathbb{R}^{n-k}$ can be embedded into $\mathbb{R}^n$ through the mapping

$$\iota : (x_1, x_2, \dots, x_{n-k}) \longmapsto (\underbrace{0, 0, \dots, 0}_{k \text{ zeros}}, x_1, x_2, \dots, x_{n-k}) .$$

Consider the set $\tilde{\mathcal{A}}$ of those diffeomorphisms $\tilde{\varphi} : V \longrightarrow U$ between open subsets of $\mathbb{R}^n$ through which the set $X \cap V$ is mapped onto the intersection $\iota(\mathbb{R}^{n-k}) \cap U$.

Put

$$\mathcal{A} = \{\varphi : \varphi \text{ has the form } \varphi = \iota^{-1} \circ \tilde{\varphi}|_{M \cap V} : X \cap V \rightarrow \mathbb{R}^{n-k}, \text{ where } \tilde{\varphi} \in \tilde{\mathcal{A}}, \text{ dom } \tilde{\varphi} = V \}.$$

Obviously, elements of $\mathcal{A}$ define a homeomorphism between open subsets of X and that of $\mathbb{R}^{n-k}$. It is also clear from the construction that the mappings $\varphi \circ \psi^{-1}$, defined on $\psi(\text{dom } \varphi \cap \text{dom } \psi)$, are smooth for any $\varphi, \psi \in \mathcal{A}$. In such a way, to prove that $\mathcal{A}$ is a $\mathcal{C}^\infty$-compatible atlas on X one has only to check that each point of X is represented in at least one chart belonging to $\mathcal{A}$.

For this purpose, consider an arbitrary point $\mathbf{x}$ of X and the gradient vectors of the coordinate functions of F at $\mathbf{x}$. Since they are linearly independent, we can obtain a basis by joining further n-k vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n-k}$ to them.

The gradient vector of the linear function $g_i(\mathbf{x}) = \langle \mathbf{e}_i, \mathbf{x} \rangle$ is $\mathbf{e}_i$ at any point, consequently, the inverse function theorem can be applied to the mapping

$$\tilde{F} = (f_1, \ldots, f_k, g_1, \ldots, g_{n-k}) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

at the point $\mathbf{x}$. According to the theorem, $\tilde{F}$ is a diffeomorphism between a neighborhood V of $\mathbf{x}$ and an open neighborhood of $\tilde{F}(\mathbf{x})$. Denote by $\tilde{\varphi}$ the restriction of $\tilde{F}$ onto V. The mapping $\tilde{\varphi}$ belongs to $\tilde{\mathcal{A}}$ and the chart $\iota^{-1} \circ (\tilde{\varphi}|_{V \cap M})$ is defined in a neighborhood of $\mathbf{x}$, so the proof is finished. ▌

Exercise. Check that the topology of $(X, \mathcal{A})$ coincides with the subspace topology inherited from $\mathbb{R}^n$, consequently, it is Hausdorff and second countable.

Definition. We say that $X \subset \mathbb{R}^n$ is an embedded (n - k) -dimensional submanifold in $\mathbb{R}^n$ if in a neighborhood U of every point $\mathbf{x} \in X$ there are functions $f_1, f_2, \ldots, f_k : U \longrightarrow \mathbb{R}$ such that the intersection of U with X is given by the equations $f_1 = f_2 = \ldots = f_k = 0$ and the vectors grad $f_1, \ldots,$ grad f_k at $\mathbf{x}$ are linearly independent.

The study of higher dimensional manifolds was launched at the beginning of the 20-th century by H. Poincaré (1854 - 1912). At that time topology was in its cradle and the abstract definition of a topological space had not been created. Poincaré worked with embedded submanifolds in $\mathbb{R}^n$. This was not a real loss of generality since by *Whitney's theorem* every differentiable manifold of dimension n is diffeomorphic to an embedded submanifold in $\mathbb{R}^{2n+1}$.

Examples.

(i) As an application, consider the set S^{n-1} of points in $\mathbb{R}^n$ the distance of which from the origin is equal to one. They are characterized by the equality

$$f(x_1, \ldots, x_n) = x_1^2 + \ldots + x_n^2 - 1 = 0.$$

The gradient vector of f at the point $\mathbf{x} \in \mathbb{R}^n$ is just $2\mathbf{x}$. It is zero only at the origin, which does not belong to S^{n-1}, so S^{n-1} is a topological manifold with a natural differentiable structure on it. S^{n-1} is called the (n-1)-dimensional sphere with the standard differentiable structure.

(ii) Most important Lie groups are obtained as closed subgroups of $Gl(n,\mathbb{R})$, defined by some equalities on the matrix entries. For example, $Gl(n,\mathbb{C})$ is isomorphic to the subgroup of $Gl(2n,\mathbb{R})$ consisting of matrices of the form $\begin{pmatrix} A & B \\ -B & A \end{pmatrix}$, where A,B are nxn matrices. The following example demonstrates how we can prove that a closed subgroup of $Gl(n,\mathbb{R})$ is a Lie group.

Consider the orthogonal group

$$O(n) = \{A \in Gl(n,\mathbb{R}) : A A^T = I\}.$$

We claim that it is an $\frac{n(n-1)}{2}$ - dimensional Lie group.

If $A = (a_{ij})$, then the equation $A A^T = I$ is equivalent to the system of equations

$$F_{ij}(A) = \sum_k a_{ik}a_{jk} = \delta_{ij}, \qquad (i,j = 1,\dots,n).$$

Since $F_{ij} = F_{ji}$, these equations are not independent. We get however an independent system of equations if we restrict ourselves to the equations with $1 \leqslant i \leqslant j \leqslant n$. The number of these equations is $\frac{n(n+1)}{2}$, so if we show the independence of the gradients then we obtain the required expression for the dimension of O(n).

The gradient of F_{ij} at $A \in O(n)$ is an nxn matrix with entries

$$\frac{\partial F_{ij}}{\partial a_{rs}} = \sum_k \left(\frac{\partial a_{ik}}{\partial a_{rs}} a_{jk} + a_{ik} \frac{\partial a_{jk}}{\partial a_{rs}} \right) = a_{js}\delta_{ir} + a_{is}\delta_{jr}$$

We show that these vectors, are orthogonal with respect to the usual scalar product on $\mathbb{R}^{n^2} = Mat(n,\mathbb{R})$. Indeed, taking the scalar product of the gradient vectors of F_{ij} and F_{kl} at A we obtain

$$\sum_{r,s} (a_{js}\delta_{ir} + a_{is}\delta_{jr})(a_{ls}\delta_{kr} + a_{ks}\delta_{lr}) =$$

$$= \sum_{r,s} a_{js}a_{ls}\delta_{ir}\delta_{kr} + a_{js}a_{ks}\delta_{ir}\delta_{lr} + a_{is}a_{ls}\delta_{jr}\delta_{kr} + a_{is}a_{ks}\delta_{jr}\delta_{lr} =$$

$$= \sum_r \delta_{jl}\delta_{ir}\delta_{kr} + \delta_{jk}\delta_{ir}\delta_{lr} + \delta_{il}\delta_{jr}\delta_{kr} + \delta_{ik}\delta_{jr}\delta_{lr} =$$

$$= \delta_{jl}\delta_{ik} + \delta_{jk}\delta_{il} + \delta_{il}\delta_{jk} + \delta_{ik}\delta_{jl} = 2\,\delta_{ik}\delta_{jl}(1+\delta_{ij}\delta_{kl}).$$

We know that the determinant of an orthogonal matrix is ±1. Thus O(n) has two components. The connected component on which determinant is 1, is the <u>special orthogonal group</u> SO(n), and it has also dimension $\frac{n(n-1)}{2}$. SO(n) is the configuration space of a rigid n-dimensional body with one fixed point. The configuration space of an n-dimensional body without a fix point is the space $SO(n) \times \mathbb{R}^n$.

For n = 2, SO(2) is the 1-dimensional group of rotations of the plane. For n = 3, SO(3) is a 3-dimensional manifold. With the help of quaternions we can show that this group is homeomorphic to the 3-dimensional projective space.

Let $\mathbb{H}$ denote the 4-dimensional space of quaternions x+yi+zj+wk, and let us identify $\mathbb{R}^3$ with the space of pure imaginary quaternions. For $0 \neq q \in \mathbb{H}$ let us denote by ρ_q the transformation $\rho_q(h)=q^{-1}hq$. Since $|q^{-1}hq|=|h|$, ρ_q is an orthogonal transformation. If h is a real number then $\rho_q(h)=h$, thus ρ_q maps $\mathbb{R} \subset \mathbb{H}$ into itself. Consequently, it maps $\mathbb{R}^3$, the orthogonal complement of $\mathbb{R}$ also into itself. The assignment $q \longmapsto \rho_q|_{\mathbb{R}^3}$ is a group homomorphism from the multiplicative group $\mathbb{H}\setminus\{\mathbf{0}\}$ to SO(3). The kernel of this homomorphism is the center of $\mathbb{H}\setminus\{\mathbf{0}\}$, i.e. $\mathbb{R}\setminus\{\mathbf{0}\}$. This homomorphism is also surjective as it follows from the following two exercises.

Exercise.

A) Show that every element of SO(3) is a rotation about a straight line.

B) Show that if $a \in \mathbb{R}^3$ is a pure imaginary quaternion, $\lambda \in \mathbb{R}$ and $q = 1+\lambda a$, then ρ_q is a rotation about a and varying λ we can obtain all rotations.

We conclude that SO(3) is isomorphic to the factorgroup $\mathbb{H}\setminus\{\mathbf{0}\}/\mathbb{R}\setminus\{\mathbf{0}\}$, but cosets of $\mathbb{R}\setminus\{\mathbf{0}\}$ in $\mathbb{H}\setminus\{\mathbf{0}\}$ are in one to one correspondence with straight lines through $\mathbf{0}$ in $\mathbb{H}$, i.e. there is a natural bijection between SO(3) and the projective space Gr(4,1).

The classification problem of n-dimensional manifolds can be formulated in different levels. For each r we may consider the category of $\mathcal{C}^r$-manifolds and r-times differentiable mappings. The following two theorems show that the classification problem of $\mathcal{C}^r$-manifolds is the same problem for all $1 \leqslant r \leqslant \omega$.

Theorem. For $r \geqslant 1$, every maximal $\mathcal{C}^r$-compatible atlas contains a $\mathcal{C}^\omega$-compatible atlas.

Theorem. It two $\mathcal{C}^\omega$-manifolds are $\mathcal{C}^1$-diffeomorphic, then they are $\mathcal{C}^\omega$-diffeomorphic.

None of the above theorems can be extended to $\mathcal{C}^0$-manifolds. There exist topological manifolds which have no $\mathcal{C}^1$-compatible atlas, and there exist homeomorphic but not diffeomorphic differentiable manifolds.

In 1956 J.W. Milnor constructed a differentiable manifold which is homeomorphic to the 7-dimensional sphere but not diffeomorphic to it. Such manifolds were given the name ”exotic spheres”. Later on an even more surprising result was published by Milnor and Kervaire. There are exactly 28 mutually non-diffeomorphic differentiable structures on a topological 7-sphere. Since then many examples of topological manifolds having many different differentiable structures have been obtained. One of the most interesting constructions is due to Donaldson, who invented exotic differentiable structures on $\mathbb{R}^4$.

We do not meet these problems in dimension two. The classification of compact surfaces is the same up to homeomorphism and diffeomorphism. The classification theorem of compact 2-dimensional manifolds gives a list of non-diffeomorphic compact surfaces and asserts that every compact surface is diffeomorphic to one of the surfaces in the list. The list of compact surfaces contains as a matter of fact two lists. The first list contains the orientable compact surfaces, the second contains the non-orientable ones.

Orientable compact surfaces. The simplest orientable closed surface is the sphere $S^2 = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| = 1\}$. The next example is the torus $T^2=S^1 x S^1$. Cutting a small disc out of the torus, we get a surface with boundary, called a handle. A typical orientable compact surface is a sphere with g handles. We can obtain this surface if we cut g holes on the surface of the sphere and glue a handle to each of them.

Non-orientable compact surfaces. The first member of this list is the real projective plane. To understand the topological structure of the projective plane, let us cut the projective plane into two parts by a hyperbola. The interior of the hyperbola has two components in the Euclidean plane, but these components are glued together along a segment of the line at infinity , so the interior is a topological disc. The exterior of the hyperbola is a long infinite band in the Euclidean plane, the "ends" of which are glued together along another segment of the line at infinity. One can see that the ends of the band are glued together by a half twist so what we get is a Möbius band. We conclude that the projective plane is the union of a disc and a Möbius band glued together along their boundaries. A typical non-orientable compact surface is a sphere with g Möbius bands. We obtain this surface cutting g discs out of the sphere and gluing to the boundary of each hole a Möbius band. Non-orientable compact surfaces can not be embedded into the 3-dimensional Euclidean space, so although one can easily construct a 3-dimensional model of a Möbius band and that of a sphere with g holes, it is impossible to glue the Möbius bands to the sphere in practice. If however we could try this in a 4-dimensional space there would be no difficulty.

Further Exercises

Exercise 8-1. The configuration space of the pentagon (closed chain of five rods in the plane) with one edge fixed is a compact surface (sometimes with singularities). What kind of surfaces can we obtain?

Exercise 8-2. Give an example of a set X with a $\mathscr{C}^0$-compatible atlas $\mathscr{A}$ on it such that the topology induced on X by $\mathscr{A}$ is (i) not Hausdorff; (ii) not second countable.

Exercise 8-3. Show that the special unitary group

$$SU(n) = \{ A \in Gl(n,\mathbb{C}) : A A^* = I, \det A = 1 \}$$

is a Lie group, determine its dimension. Prove that SU(2) is diffeomorphic to the 3-dimensional sphere S^3.

Exercise 8-4. Which surface shall we get from the classification list if we glue to the sphere $k \geqslant 1$ Möbius bands and l handles?

Exercise 8-5. Let P be a complex polynomial of degree k having k different roots. Consider the subset of $\mathbb{C}^2$ defined by

$$M = \{(z,w) \in \mathbb{C}^2 : z^l = P(w)\}.$$

Show that M is diffeomorphic to a sphere with g handles with N points omitted. Express g and N in terms of k and l.

Unit 9. The Tangent Bundle

The tangent space of a submanifold of $\mathbb{R}^n$, identification of tangent vectors with derivations at a point, the abstract definition of tangent vectors, the tangent bundle; the derivative of a smooth map.

The aim of this chapter is to give and reconcile different commonly used definitions of a tangent vector to a manifold. Before passing over to the abstract situation, we shall deal with submanifolds of $\mathbb{R}^n$.

Definition. Smooth mappings $\gamma : [a,b] \longrightarrow M$ of an interval into a differentiable manifold $(M, \mathcal{A})$ are called smooth curves in the manifold.

Definition **A**. Let M be a differentiable manifold embedded in $\mathbb{R}^n$, $\mathbf{x}_0 \in M$. A vector $\mathbf{v}$ is called a tangent vector to M at $\mathbf{x}_0$ if there is a smooth curve $\mathbf{x}: [-\varepsilon, \varepsilon] \longrightarrow M$ passing through $\mathbf{x}_0 = \mathbf{x}(0)$ such that $\mathbf{v} = \mathbf{x}'(0)$. The tangent space $T_{\mathbf{x}_0}M$ of M at $\mathbf{x}_0$ is the set of all tangent vectors to M at $\mathbf{x}_0$.

Theorem. Let us suppose that a k-dimensional manifold M embedded in $\mathbb{R}^n$ is given in a neighborhood U of $\mathbf{x}_0 \in M$ by a system of equalities $f_1 = \ldots = f_{n-k} = 0$, where $f_1, \ldots, f_{n-k}$ are smooth functions on U such that the vectors grad $f_1(\mathbf{x}_0), \ldots,$ grad $f_{n-k}(\mathbf{x}_0)$ are linearly independent at $\mathbf{x}_0$. Then the tangent space of M at $\mathbf{x}_0$ consists of the vectors orthogonal to grad $f_1(\mathbf{x}_0), \ldots,$ grad $f_{n-k}(\mathbf{x}_0)$.

Corollary. The tangent space of a k-dimensional submanifold of $\mathbb{R}^n$ is a k-dimensional linear subspace of $\mathbb{R}^n$.

Proof. If $\mathbf{x}: [-\varepsilon, \varepsilon] \longrightarrow \mathbb{R}^n$ is a smooth curve having coordinate functions $x_1, \ldots, x_n$ and lying on M, then we have

$$f_i(x_1(t), \ldots, x_n(t)) = 0 \qquad i = 1, \ldots, n-k$$

for each $t \in [-\varepsilon, \varepsilon]$. Differentiating by t we get

$$\frac{\partial f_i}{\partial x_1}(\mathbf{x}(0))x_1'(0) + \ldots + \frac{\partial f_i}{\partial x_n}(\mathbf{x}(0))x_n'(0) = 0$$

which means that the vectors grad $f_i(\mathbf{x}(0))$ and $\mathbf{x}'(0)$ are orthogonal.

Now let us prove that if a vector $\mathbf{v}$ is orthogonal to the vectors grad $f_i(\mathbf{x}_0)$, $1 \leqslant i \leqslant n-k$, then $\mathbf{v}$ is a tangent vector.

Let us take a smooth local parameterization $F: \mathbb{R}^k \longrightarrow M \subset \mathbb{R}^n$ of M around the point $\mathbf{x}_0$. The curve $t \longmapsto F(F^{-1}(\mathbf{x}_0) + t\mathbf{y})$, where $\mathbf{y} \in \mathbb{R}^k$ fixed, is a curve on M passing through $\mathbf{x}_0$. The speed vector of this curve for $t = 0$ is

$$\frac{\partial F}{\partial x_1}(F^{-1}(\mathbf{x}_0))y_1 + \ldots + \frac{\partial F}{\partial x_k}(F^{-1}(\mathbf{x}_0))y_k,$$

where $y_1,\ldots,y_k$ are the coordinates of $\mathbf{y}$. By the construction of local parameterizations of embedded manifolds, F is a restriction of a diffeomorphism between open subsets of $\mathbb{R}^n$ onto $\mathbb{R}^k$, consequently, the vectors $\frac{\partial F}{\partial x_i}(F^{-1}(\mathbf{x}_0))$ are linearly independent. We conclude, that the tangent space is contained in the k-dimensional linear subspace orthogonal to grad $f_1(\mathbf{x}_0),\ldots,\text{grad } f_{n-k}(\mathbf{x}_0)$ and contains the k-dimensional linear subspace spanned by the vectors $\frac{\partial F}{\partial x_i}(F^{-1}(\mathbf{x}_0))$ $1 \leqslant i \leqslant k$, which means that both linear subspaces coincide with the tangent space.▌

The definition of tangent vectors can also be given in intrinsic terms, independent of the embedding of M into $\mathbb{R}^n$.

Let us define an equivalence relation on the set

$$\mathscr{C}urve(M,p) = \{\gamma\colon[-\varepsilon,\varepsilon]\longrightarrow M : \gamma(0) = p\},$$

consisting of curves passing through $p \in M$, by calling two curves $\gamma_1,\gamma_2 \in \mathscr{C}urve(M,p)$ equivalent if $(\mathbf{x}\circ\gamma_1)'(0) = (\mathbf{x}\circ\gamma_2)'(0)$ for some chart $\mathbf{x}$ around p. Then this condition is true for any chart (prove this!).

<u>Definition</u> **B**. A <u>tangent vector</u> to a manifold M at the point $p \in M$ is an equivalence class of curves belonging to $\mathscr{C}urve(M,p)$. The set of equivalence classes is called the <u>tangent space of M at p</u> and denoted by T_pM.

Given a chart $\mathbf{x}$ around p, we can establish a one-to-one correspondence between the equivalence classes and points of $\mathbb{R}^m$, ($m = \dim M$), assigning to the equivalence class of a curve $\gamma \in \mathscr{C}urve(M,p)$ the vector $(\mathbf{x}\circ\gamma)'(0) \in \mathbb{R}^m$. With the help of this identification, we can introduce a vector space structure on the tangent space, not depending on the choice of the chart.

For embedded manifolds definition **B** agrees with definition **A**. The advantage of definition **B** lies in the fact that it is applicable also for abstract manifolds, not embedded anywhere.

<u>Definition</u>. If $\mathbf{x} = (x_1,\ldots,x_m)$ is a chart on the manifold M around the point p, $\gamma \in \mathscr{C}urve(M,p)$, then the numbers $(x_1\circ\gamma)'(0)$,..., $(x_m\circ\gamma)'(0)$ are called the <u>components of the tangent vector</u> represented by γ <u>with respect to the chart</u> $\mathbf{x}$.

The main difficulty of defining tangent vectors to a manifold is due to the fact that an abstract manifold might not be embedded into a fixed finite dimensional linear space. Nevertheless, there is a universal embedding of each differentiable manifold into an infinite dimensional linear space.

Let us denote by $\mathcal{F}$ (M) the linear vector space of smooth functions on M, and by $\mathcal{F}^*$ (M) the dual space of $\mathcal{F}$ (M) that is the space of linear functions on $\mathcal{F}$ (M), and consider the embedding ι of M into $\mathcal{F}^*$ (M) defined by the formula

$$[\iota(p)](f) = f(p), \text{ where } p \in M,\ f \in \mathcal{F}(M).$$

Having embedded the manifold M into $\mathcal{F}^*$ (M), we can define tangent vectors to M to be elements of the linear space $\mathcal{F}^*$ (M).

<u>Definition</u>. Let M be a differentiable manifold, p ∈ M. We say that a linear function $D \in \mathcal{F}^*$ (M) defined on smooth functions on M is a <u>derivation at the point p</u> if the equality

$$D(fg) = D(f)g(p) + f(p)D(g)$$

holds for every $f,g \in \mathcal{F}$ (M).

Each curve $\gamma \in \mathcal{C}urve(M,p)$ defines a derivation at the point p by the formula $D_{\gamma'(0)}(f) = (f\circ\gamma)'(0)$, where $f \in \mathcal{F}$ (M). $D_{\gamma'(0)}$ is the speed vector of the curve $\iota\circ\gamma$ in $\mathcal{F}^*$(M). Since two curves define the same derivation iff they are equivalent, there is a one-to-one correspondence between the equivalence classes of curves and the derivations obtained as $D_{\gamma'(0)}$ for some γ.

<u>Definition</u> **C**. A <u>tangent vector</u> to a manifold M at the point p ∈ M is a derivation of the form $D_{\gamma'(0)}$, where $\gamma \in \mathcal{C}urve(M,p)$.
The <u>tangent space</u> T_pM of M at the point p is the set of derivations $D_{\gamma'(0)}$ along curves in M passing through $p = \gamma(0)$.

<u>Theorem</u>. The tangent space to a differentiable manifold M at the point p ∈ M coincides with the space of derivations on $\mathcal{F}$ (m) at p, which is a linear space having the same dimension as M has.

<u>Lemma</u> 1. If $f \in \mathcal{F}$ (M) is a constant function and D is a derivation at a point p ∈ M, then D(f) = 0.

<u>Proof</u>. Because of linearity, it is enough to show that D(**1**)=0, where **1** is the constant 1 function on M. But we have

$$D(\mathbf{1}) = D(\mathbf{1}\ \mathbf{1}) = D(\mathbf{1})\ \mathbf{1}(p) + \mathbf{1}(p)\ D(\mathbf{1}) = 2\ D(\mathbf{1}).\ \blacksquare$$

Lemma 2. If two functions $f,g \in M$ coincide on a neighborhood U of $p \in M$ and D is a derivation at p then $D(f) = D(g)$.

Proof.

Sublemma. If $\mathbf{x} \in \mathbb{R}^n$ and $B(\mathbf{x},\varepsilon)$ is a fixed open ball about it, then there exists a smooth function $h:\mathbb{R}^n \longrightarrow \mathbb{R}$ such that $h(\mathbf{y})$ is equal to 1 if $\mathbf{y} \in B(\mathbf{x},\varepsilon/2)$, positive if $\mathbf{y} \in B(\mathbf{x},\varepsilon)$ and zero if $\mathbf{y} \notin B(\mathbf{x},\varepsilon)$.

Define the function h_0 of the real variable t by the formula

$$h_0(t) = \begin{cases} e^{-(1-t^2)^{-1}} & \text{if } t \in (-1,1) \\ 0 & \text{otherwise.} \end{cases}$$

It is a good exercise to prove that h_0 is a smooth function on $\mathbb{R}$. Set $h_1(\mathbf{y}) := h_0(4\|\mathbf{y}\| /\varepsilon)$, let χ denote the characteristic function of the ball $B(\mathbf{x},3\varepsilon/4)$, and define the function h_2 as follows

$$h_2(\mathbf{y}) := \int_{\mathbb{R}^n} \chi(\mathbf{z})h_1(\mathbf{y}-\mathbf{z})d\mathbf{z}.$$

If we put $h(\mathbf{y}) = h_2(\mathbf{y})/h_2(\mathbf{x})$ then we get a desirable function.

Now let us prove the lemma. Using the construction above we can define a smooth function h on M which is zero outside U and such that $h(p) = 1$. In this case $h(f-g)$ is the constant 0 function on M. Thus we have

$$0 = D(\mathbf{0}) = D(h(f-g)) = D(h)\ (f(p)-g(p)) + h(p)\ D(f-g) = D(f) - D(g). \blacksquare$$

Remarks.

i) The sublemma shows that the mapping $\iota:M\longrightarrow \mathcal{F}^*(M)$ above is indeed an inclusion. If $p \neq q$ are distinct points of M, then there is a smooth function h on M such that

$$[\iota(p)](h) = h(p) = 1 \neq [\iota(q)](h) = h(q) = 0.$$

ii) We can extend a derivation D at a point p on functions f defined only in a neighborhood U of p by taking a smooth function h on M such that h is zero outside U and constant 1 in a neighborhood of p and putting $D(f) := D(\tilde{f})$, where

$$\tilde{f}(x) = \begin{cases} f(x)h(x) & \text{for } x \in U \\ 0 & \text{for } x \notin U. \end{cases}$$

By lemma 2 this extension of D is correctly defined.

Lemma 3. Let $f:B\longrightarrow \mathbb{R}$ be a smooth function defined on an open ball $B \subset \mathbb{R}^n$ around the origin. Then there exist smooth functions g_i $1\leqslant i\leqslant n$ on B such that

$$f(\mathbf{x}) = f(\mathbf{0}) + \sum_{i=1}^{n} x_i g_i(\mathbf{x}) \qquad \text{for } \mathbf{x} = (x_1,\ldots,x_n) \in B$$

and

$$g_i(\mathbf{0}) = \frac{\partial f}{\partial x_i}(\mathbf{0}).$$

Proof. Since

$$f(\mathbf{x}) - f(\mathbf{0}) = \int_0^1 \frac{d\, f(t\mathbf{x})}{dt} dt = \int_0^1 \sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(t\mathbf{x})\, dt =$$

$$= \sum_{i=1}^n x_i \int_0^1 \frac{\partial f}{\partial x_i}(t\mathbf{x})\, dt, \quad \text{we may take } g_i(\mathbf{x}) = \int_0^1 \frac{\partial f}{\partial x_i}(t\mathbf{x})\, dt. \blacksquare$$

Now we are ready to prove the theorem.

Let us take a differentiable manifold $(M,\mathcal{A})$ and a chart $\mathbf{x} = (x_1,\ldots,x_n) \in \mathcal{A}$ defined in a neighborhood of $p \in M$.

Define the derivations $\partial_i(p)$ as follows

$$\Big[\partial_i(p)\Big](f) := \frac{\partial\, f\circ\mathbf{x}^{-1}}{\partial x_i} (\mathbf{x}(p)).$$

We prove that the derivations $\partial_i(p)$ form a basis in the space of derivations at p. They are linearly independent since if we have

$$\sum_{i=1}^n \alpha_i \partial_i(p) = 0,$$

then applying this derivation to the j-th coordinate function x_j we get

$$\sum_{i=1}^n \alpha_i \frac{\partial x_j}{\partial x_i}(\mathbf{x}(p)) = \alpha_j = 0.$$

On the other hand, if D is an arbitrary derivation at p, then we have

$$D = \sum_{i=1}^n D(x_i)\partial_i(p).$$

Indeed, let $f \in \mathcal{F}(M)$ be an arbitrary smooth function on M and apply lemma 3 to $f\circ\mathbf{x}^{-1}$ around $\mathbf{x}(p)$. We obtain functions g_i defined around $\mathbf{x}(p)$ such that

$$f = f(p) + \sum_{i=1}^n (x_i - x_i(p)) g_i\circ\mathbf{x} \quad \text{and} \quad g_i(\mathbf{x}(p)) = \frac{\partial\, f\circ\mathbf{x}^{-1}}{\partial x_i} (\mathbf{x}(p)) .$$

In this case however we have

$$D(f) = D(f(p)) + \sum_{i=1}^n D((x_i - x_i(p)))\ g_i(\mathbf{x}(p)) + (x_i(p) - x_i(p))\ D(g_i\circ\mathbf{x}) =$$

$$\sum_{i=1}^n D(x_i) \frac{\partial\, f\circ\mathbf{x}^{-1}}{\partial x_i} (\mathbf{x}(p)) = \sum_{i=1}^n D(x_i) \Big[\partial_i(p)\Big](f).$$

To finish the proof, we only have to show that every derivation at the point p can be obtained as a speed vector of a curve passing through p.

Define the curve $\gamma : [-\varepsilon,\varepsilon] \longrightarrow M$ by the formula

$$\gamma(t) := \mathbf{x}^{-1}(\mathbf{x}(p) + (t\alpha_1,\ldots,t\alpha_n)).$$

Then obviously the speed vector $\gamma'(0)$ is just $\sum_{i=1}^n \alpha_i \frac{\partial}{\partial x_i}(p)$. ■

The tangent bundle

The union of the tangent spaces of M at the various points, $\bigcup_{p\in M} T_pM$, has a natural differentiable manifold structure, the dimension of which is twice the dimension of M.

This manifold is called the tangent bundle of M and is denoted by TM. A point of this manifold is a vector D, tangent to M at some point p. Local coordinates on TM are constructed as follows. Let $\mathbf{x} = (x_1,\dots,x_n)$ be a chart on M the domain U of which contains p, and $D(x_1),\dots,D(x_n)$ be the components of D in the basis $\partial_i(p)$. Then the mapping

$$D \longmapsto (x_1(p),\dots,x_n(p),D(x_1),\dots,D(x_n))$$

give a local coordinate system on $\bigcup_{p\in U} T_pM \subset TM$. The set of all local coordinate systems constructed this way forms a $\mathcal{C}^\infty$-compatible atlas on TM, that turns TM into a differentiable manifold.

Exercise. Check the last statement.

The mapping $\pi : TM \longrightarrow M$ which takes a tangent vector D to the point $p \in M$ at which the vector is tangent to M is called the natural projection. The inverse image of a point $p \in M$ under the natural projection is the tangent space T_pM. This space is called the fiber of the tangent bundle over the point p.

The derivative of a map

Definition. Let $f:M \longrightarrow N$ be a smooth mapping between the differentiable manifolds $(M,\mathcal{A})$, $(N,\mathcal{B})$, and let $p \in M$. The derivative of f at the point p is the linear map of the tangent spaces $f'_p : T_pM \longrightarrow T_{f(p)}N$, which is given in the following way.

Let $D \in T_pM$ and consider a curve $\gamma : [-\varepsilon,\varepsilon] \longrightarrow M$ with $\gamma(0)= p$, and speed vector D. Then $f'_p(D)$ is the tangent vector represented by the curve $f\circ\gamma$.

Proposition. The derivative f'_p is correctly defined (does not depend on the choice of γ) and is linear.

Proof. We derive a formula for f'_p using local coordinates which will show both parts of the proposition clearly.

Let $\mathbf{x} = (x_1,\dots,x_m)$ and $\mathbf{y} = (y_1,\dots,y_n)$ be local coordinates in a neighborhood of $p \in M$ and $f(p) \in N$ respectively.

If the components of D in the basis $\partial_i(p)$ corresponding to the chart $\mathbf{x}$ are $\{ \alpha_i : 1 \leqslant i \leqslant m \}$ then we have $(x_i \circ \gamma)'(0) = \alpha_i$. Observe, that α_i depends only on D. The components $\{ \beta_j : 1 \leqslant j \leqslant n \}$ of $f'_p(D)$ in the basis $\tilde{\partial}_j(f(p))$ generated by the chart $\mathbf{y}$ can be computed by the formula $\beta_j = (y_j \circ f \circ \gamma)'(0)$. Denote by $\tilde{f}_j$ the j-th coordinate function of the mapping $\mathbf{y} \circ f \circ \mathbf{x}^{-1}$, i.e.

$$\tilde{f}_j = y_j \circ f \circ \mathbf{x}^{-1}.$$

Then we have

$$\beta_j = (y_j \circ f \circ \gamma)'(0) = [(y_j \circ f \circ \mathbf{x}^{-1}) \circ \mathbf{x} \circ \gamma]'(0) = [\tilde{f}_j \circ (\mathbf{x} \circ \gamma)]'(0) =$$

$$= \sum_{i=1}^{m} \frac{\partial \tilde{f}_j}{\partial x_i} (\mathbf{x}(p))(x_i \circ \gamma)'(0) = \sum_{i=1}^{m} \frac{\partial \tilde{f}_j}{\partial x_i} (\mathbf{x}(p))\alpha_i,$$

which shows that $f'_p(D)$ depends only on D and that f'_p is a linear mapping the matrix of which in the bases $\left\{\partial_i(p)\right\}$ and $\left\{\tilde{\partial}_j(f(p))\right\}$ is

$$\left(\frac{\partial \tilde{f}_j}{\partial x_i} (\mathbf{x}(p)) \right)_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}. \blacksquare$$

Unit 10. The Lie Algebra of Vector Fields

Vector fields and ordinary differential equations; basic results of the theory of ordinary differential equations (without proof); the Lie algebra of vector fields and the geometric meaning of Lie bracket, commuting vector fields, Lie algebra of a Lie group.

Definition. A smooth vector field X over a differentiable manifold M is a smooth mapping of M into its tangent bundle, such that $X(p) \in T_pM$ for each $p \in M$.

Obviously, smooth vector fields over M form a real vector space with respect to the operations

$$(X + Y)(p) := X(p) + Y(p), \quad (\lambda X)(p) := \lambda X(p),$$

where X, Y are vector fields, $\lambda \in \mathbb{R}$, $p \in M$. We can multiply vector fields by smooth functions as well, by the rule

$$(f X)(p) := f(p) X(p).$$

We denote by $\mathfrak{X}(M)$ the vector space of smooth vector fields.

Associated to a local coordinate system $\mathbf{x} = (x_1,\dots,x_n)$ on M, there is a basis of T_pM at each $p \in \mathrm{dom}\,\mathbf{x}$, formed by the tangent vectors $\left\{\partial_i(p):\ 1\leqslant i\leqslant n\right\}$. The mapping $\partial_i : p \longmapsto \partial_i(p)$ gives a local smooth vector field in the domain of the chart for each i. Thus, every smooth vector field X can be written in the form

$$X = \sum_{i=1}^{n} X_i\, \partial_i,$$

where the X_i-s are smooth functions on the domain of $\mathbf{x}$. The functions X_i are called the components of the vector field X.

Given a smooth vector field X on a manifold M, we may pose the following problem. Find those smooth curves $\gamma : (a,b) \longrightarrow M$ on M for which the speed of γ at $t \in (a,b)$ is $X(\gamma(t))$. Such curves are called the integral curves of the vector field. Obviously, a restriction of an integral curve onto a subinterval is also an integral curve, therefore, it is enough to look for the maximal integral curves which can not be extended to an integral curve defined on a larger interval.

Trying to solve the problem, we find that it reduces to an ordinary differential equation of first order. Indeed, γ is an integral curve if and only if for each chart $\mathbf{x} = (x_1,\dots,x_n)$, the "vector-valued" function

$$\mathbf{f} = \mathbf{x}\circ\gamma : (a,b) \longrightarrow \mathbb{R}^n$$

satisfies the differential equation

$$\mathbf{f}'(t) = (X_1\circ\mathbf{x}^{-1}\circ\mathbf{f}(t),\dots,X_n\circ\mathbf{x}^{-1}\circ\mathbf{f}(t)).$$

Actually, finding integral curves of a vector field is the same problem as solving an ordinary differential equation, only the language of formulation is different. Translating the basic results of the theory of ordinary differential equations into the language of geometry we get the following theorems, we mention without proof.

Theorem.

i) (Existence and uniqueness of solutions). Let X be a smooth vector field on a differentiable manifold M. Then for each point $p \in M$ there exists a unique maximal integral curve $\gamma_p: (a,b) \longrightarrow M$ of the vector field X such that $0 \in (a,b)$ and $\gamma_p(0) = p$ (a and b might be $-\infty$ and ∞ respectively).

ii) (straightening vector fields). Let X be an arbitrary vector field on a manifold M and $p \in M$ such that $X(p)\neq 0$. Then there exists a chart $\mathbf{x} =(x_1,\dots,x_n)$ around p for which $X = \partial_1$. This means that the derivative of the mapping $\mathbf{x}$ turns the vector field X into a constant vector field on $\mathbb{R}^n$.

iii) (Unboundedness of solutions). If a (or b) finite then no compact subset of M contains the image $\gamma_p((a,0))$ (or $\gamma_p((0,b))$).

iv) (differentiable dependence on the initial point). Let us define the set $U_t \subset M$ for $t \in \mathbb{R}$ as follows

$$U_t = \{p \in M : t \in \operatorname{dom} \gamma_p\}.$$

Then U_t is an open subset of M and the mapping $H_t:U_t \longrightarrow M$ defined by $H_t(p) = \gamma_p(t)$ is a diffeomorphism between U_t and U_{-t}. If, furthermore, the expression $H_{t_1}(H_{t_2}(p))$ is defined, then so is $H_{t_1+t_2}(p)$ and $H_{t_1}(H_{t_2}(p)) = H_{t_1+t_2}(p)$. The family $\{H_t : t \in \mathbb{R}\}$ is called the one-parameter family of diffeomorphisms or the flow generated by the vector field.

THE LIE ALGEBRA OF VECTOR FIELDS

Since tangent vectors to a manifold at a point are identified with derivations at the point, vector fields can be considered to be differential operators assigning to a smooth function another smooth function by the formula

$$[X\ (f)](p) = [X(p)](f), \text{ where } X \in \mathfrak{X}(M),\ f \in \mathcal{F}(M),\ p \in M.$$

In this sense, a vector field X is *a linear mapping* $X:\mathcal{F}(M)\longrightarrow \mathcal{F}(M)$, *satisfying the "Leibniz' rule"*

$$X(fg) = X(f)g + fX(g).$$

Definition. Let A and B be two linear endomorphisms of a vector space V. Then the linear mapping $[A,B] = A \circ B - B \circ A$ is called the commutator of them.

Proposition. The commutator of linear mappings satisfies the following identities ($A,B,C \in \mathcal{E}nd((V)$, $\lambda \in \mathbb{R}$)

i) $[A+B,C] = [A,C]+[B,C] \quad [C,A+B] = [C,A]+[C,B]$

$[\lambda A,B] = [A,\lambda B] = \lambda[A,B]$ (bilinearity)

ii) $[A,B] = -[B,A]$ (anti-commutation)

iii) $[A,[B,C]]+[B,[C,A]]+[C,[A,B]]= 0$ (Jacobi identity)

Proof. We prove only iii), the rest is left to the reader.

$$\begin{aligned}
&[A,[B,C]]+[B,[C,A]]+[C,[A,B]] = \\
&= [A,(B\circ C-C\circ B)]+[B,(C\circ A-A\circ C)]+[C,(A\circ B-B\circ A)] = \\
&= A\circ(B\circ C-C\circ B)-(B\circ C-C\circ B)\circ A + B\circ(C\circ A-A\circ C) - \\
&\quad -(C\circ A-A\circ C)\circ B + C\circ(A\circ B-B\circ A)-(A\circ B-B\circ A)\circ C = \\
&= A\circ B\circ C - A\circ C\circ B - B\circ C\circ A + C\circ B\circ A + \\
&\quad +B\circ C\circ A - B\circ A\circ C - C\circ A\circ B + A\circ C\circ B + \\
&\quad +C\circ A\circ B - C\circ B\circ A - A\circ B\circ C + B\circ A\circ C = \\
&= 0 \ \blacksquare
\end{aligned}$$

Definition. Let us suppose that a linear space L is endowed with a bilinear mapping $[\ ,\ ]:L\text{x}L\longrightarrow L$ satisfying conditions i), ii), and iii) of the above proposition. Then the pair $(L\ ,[,])$ is called a Lie algebra.

Proposition. Let X and Y be two smooth vector fields on a manifold M. Considering them to be linear endomorphisms of the vector space of smooth functions $\mathcal{F}(M)$, the commutator $[X,Y]$ of them is also a vector field.

Proof. The commutator $[X,Y]$ is a linear endomorphism of $\mathcal{F}(M)$ so we only have to check that it satisfies the Leibniz' rule. For $f,g \in \mathcal{F}(M)$ we have

$$[X,Y](fg) = (X\circ Y-Y\circ X)(fg) = X(Y(fg))-Y(X(fg)) =$$

$$= X\ (\ Y(f)g + fY(g)\ ) + Y(\ X(f)g + fX(g)\ )$$

$$= X\circ Y(f)g + Y(f)X(g) + X(f)Y(g) + fX\circ Y(g) - Y\circ X(f)g - X(f)Y(g) -$$

$$- Y(f)X(g) - fY\circ X(g) =$$

$$= [X,Y](f)g + f[X,Y](g). \blacksquare$$

Corollary. The commutator of vector fields, which is generally called the Lie bracket of them, is a binary operation on $\mathfrak{X}(M)$, giving the space of vector fields a Lie algebra structure.

Proposition. Let us choose a local coordinate system $(x_1,\dots,x_n)$ on M and denote by $\partial_1,\dots,\partial_n$ the associated coordinate vector fields. Then we have

i) $[\partial_i,\partial_j] = 0$;

ii) $[fX,gY] = fg[X,Y] + fX(g)Y - gY(f)X$ for each $X,Y \in \mathfrak{X}(M)$, $f,g \in \mathcal{F}(M)$;

iii) if $X = \sum_{i=1}^{n} f_i\partial_i$, $Y = \sum_{i=1}^{n} g_i\partial_i$ are arbitrary vector fields,

then

$$[X,Y] = \sum_{i=1}^{n} (X(g_i)-Y(f_i))\partial_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} f_j \frac{\partial g_i}{\partial x_j} - g_j \frac{\partial f_i}{\partial x_j}\right)\partial_i .$$

Proof. i) The first part of the proposition is equivalent to Young's theorem, (known from multivariable calculus), which says that for any smooth function, defined on an open subset of $\mathbb{R}^n$, the mixed partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$ and $\frac{\partial^2 f}{\partial x_j \partial x_i}$ are equal.

ii) Let h be an arbitrary smooth function on M, and apply the operator $[fX,gY]$ to it.

$$[fX,gY](h) = fX(g\,Y(h))-gY(fX(h)) = fX(g)Y(h)+fgX(Y(h))- gY(f)X(h)- gfY(X(h))$$
$$= \Big(fg[X,Y] + fX(g)Y - gY(f)X\Big)(h).$$

iii) Using i) and ii) we get

$$[X,Y] = \left[\sum_{i=1}^{n} f_i\partial_i , \sum_{j=1}^{n} g_j\partial_j \right] = \sum_{i=1}^{n} \sum_{j=1}^{n} [f_i\partial_i, g_j\partial_j] =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} f_i\partial_i(g_j)\partial_j - g_j\partial_j(f_i)\partial_i = \sum_{i=1}^{n} (X(g_i) - Y(f_i))\partial_i . \blacksquare$$

Suppose that we are given two vector fields X and Y on an open subset of $\mathbb{R}^n$. The corresponding flows H_s and G_t do not commute in general: $H_s\circ G_t \neq G_t\circ H_s$.

To measure the lack of commutation of the flows H_s and G_t, we consider the difference

$$\Phi(s,t;p) = G_t\circ H_s(p) - H_s\circ G_t(p), \text{ for a fixed point } p.$$

Φ is a differentiable function of s and t and it is **0** if t or s is zero. This means, that in the Taylor expansion of Φ around $(0,0;p)$

$$\Phi(s,t;p) = \Phi(0,0;p)+\left(s\,\frac{\partial\,\Phi}{\partial\,s}\,(0,0;p) + t\,\frac{\partial\,\Phi}{\partial\,t}\,(0,0;p)\right) +$$

$$+\left(\frac{s^2}{2}\,\frac{\partial^2\Phi}{\partial\,s^2}\,(0,0;p) + st\,\frac{\partial^2\Phi}{\partial s\partial t}\,(0,0;p) + \frac{t^2}{2}\,\frac{\partial^2\Phi}{\partial\,t^2}\,(0,0;p)\right) + o(s^2+t^2)$$

the only non-zero partial derivative is $\frac{\partial^2\Phi}{\partial s\partial t}\,(0,0;p)$.

<u>Claim</u>. Through the natural identification of the tangent space of $\mathbb{R}^n$ at p with the vectors of $\mathbb{R}^n$, the vector $\frac{\partial^2\Phi}{\partial s\partial t}\,(0,0;p)$ corresponds to the tangent vector $[X,Y](p)$.

<u>Proof</u>. Put $X = \sum_{i=1}^{n} f_i\partial_i$, $Y = \sum_{i=1}^{n} g_i\partial_i$, where ∂_i denotes the vector field $\frac{\partial}{\partial x_i}$. Let us compute first the vector $\frac{\partial^2}{\partial s\partial t}\,H_s\circ G_t(p)$ at $s = t = 0$.

We have $\left(\frac{\partial}{\partial\,s}\,H_s\circ G_t(p)\right)(0,t) = \left(\frac{\partial}{\partial\,s}\,\gamma_{G_t(p)}(s)\right)(0,t) = X(G_t(p))$.

Differentiating by t,

$$\frac{\partial^2}{\partial s\partial t}\,H_s\circ G_t(p)\Big|_{s=t=0} = \frac{d}{d\,t}\,X(G_t(p))\Big|_{t=0} = \sum_{i=1}^{n}\left(Y(f_i)\partial_i\right)(p).$$

A similar computation shows that

$$\frac{\partial^2}{\partial s\partial t}\,G_t\circ H_s(p)\Big|_{s=t=0} = \sum_{i=1}^{n}\left(X(g_i)\partial_i\right)(p).$$

Subtracting these equalities we get

$$\frac{\partial^2\Phi}{\partial s\partial t}\;(0,0;p) = \left(\sum_{i=1}^{n}(X(g_i) - Y(f_i))\partial_i\right)(p) = [X,Y](p). \blacksquare$$

Now returning to the Taylor expansion of Φ, we see that

$$\Phi(s,t;p) = st\,\frac{\partial^2\Phi}{\partial s\partial t}\,(0,0;p) + o(s^2+t^2) = st[X,Y](p) + o(s^2+t^2)$$

In particular, we obtain the following expression for $[X,Y](p)$.

$$\boxed{[X,Y](p) = \lim_{t\to 0}\,(G_t\circ H_t(p) - H_t\circ G_t(p))/t^2.}$$

<u>Definition</u>. We say that two vector fields are <u>commuting</u> if their Lie bracket is the zero vector field.

<u>Theorem</u>. Let $\{H_t : t \in \mathbb{R}\}$ and $\{G_t : t \in \mathbb{R}\}$ be the one-parameter families

of diffeomorphisms generated by the vector fields X and Y respectively and suppose that the vector fields X and Y are commuting. Then the diffeomorphisms H_s and G_t are commuting as well in the following sense.

For each point p of the manifold there exists a positive ε (depending on p) such that for any pair of real numbers s,t satisfying the inequality $|s| + |t| \leqslant \varepsilon$ the expressions $H_s(G_t(p))$ and $G_t(H_s(p))$ are defined and coincide: $H_s(G_t(p)) = G_t(H_s(p))$.

Proof. If both X and Y vanishes at p then $H_s(p) = G_t(p) = p$ for any s and t and thus the assertion holds trivially. We may thus suppose that one of the vectors X(p), Y(p), say X(p) is not zero. By the theorem on straightening vector fields we may suppose that the manifold is an open subset of $\mathbb{R}^n$, with coordinates $(x_1,\dots,x_n)$, and the vector field X coincides with the basis vector field ∂_1 .Let $Y = \sum_{i=1}^{n} g_i\partial_i$ be the splitting of Y into a linear combination of the basis vector fields ∂_i. By the formula for the Lie bracket of vector fields we have

$$0 = [X,Y] = \sum_{i=1}^{n} \frac{\partial g_i}{\partial x_1} \partial_i \implies \frac{\partial g_i}{\partial x_1} = 0 \quad \text{for each } i.$$

Consequently, the functions g_i do not depend on x_1, thus the vector field Y is invariant under translations parallel to the vector $\mathbf{e}_1=(1,0,\dots,0)$. This implies that if γ is an integral curve of the vector field Y then so is $\gamma + t\mathbf{e}_1$ for any t (the domain of $\gamma + t\mathbf{e}_1$ is an open subset of the domain of γ). On the other hand, the diffeomorphism H_t is just a translation by the vector $t\mathbf{e}_1$. So, for small s and t, we have

$$G_t(H_s(p)) = G_t(p + s\mathbf{e}_1) = \gamma_{p + s\mathbf{e}_1}(t) = (\gamma_p + s\mathbf{e}_1)(t) = \gamma_p(t) + s\mathbf{e}_1 =$$

$$= G_t(p) + s\mathbf{e}_1 = H_s(G_t(p)). \blacksquare$$

THE LIE ALGEBRA OF A LIE GROUP

Let $F:M \longrightarrow N$ be a diffeomorphism between two manifolds. F' defines a bijection between $\mathfrak{X}(M)$ and $\mathfrak{X}(N)$, and this bijection is a Lie algebra isomorphism.

Let G be a Lie group, $g \in G$. Denote by L_g the left translation by g, i.e., $L_g:G \longrightarrow G$, $L_g(h) = gh$. L_g is a diffeomorphism, its inverse is $L_{(g^{-1})}$. A vector field $X \in \mathfrak{X}(G)$ is called left invariant if $L_g'(X) = X$ for all $g \in G$. Since $L_g'(X) = X$ and $L_g'(Y) = Y$ implies $L_g'[X,Y] = [L_g'(X),L_g'(Y)] = [X,Y]$, left invariant vector fields form a Lie subalgebra of $\mathfrak{X}(G)$.

Definition. The Lie algebra of left invariant vector fields of a Lie group is called the Lie algebra of the Lie group.

If $X \in \mathfrak{X}(G)$ is left invariant, then $X(g)=L_g'(X(e))$, thus, a left invariant vector field is uniquely determined by the vector $X(e) \in T_eG$. (e is the unit element of the group G.) Since every vector in T_eG extends to a left invariant vector field this way, the assignment $X \longmapsto X(e)$ yields a linear isomorphism between the vector space of left invariant vector fields on G and T_eG. As a consequence, we obtain that *the Lie algebra of a Lie group is finite dimensional, its dimension is the same as that of the Lie group.*

As an example, let us determine the Lie algebra of $Gl(n,\mathbb{R})$. $Gl(n,\mathbb{R})$ is an open subset in $Mat(n,\mathbb{R}) \cong \mathbb{R}^{n^2}$, so its manifold structure is given by one chart, the embedding. Tangent spaces at different points can be identified with the linear space $Mat(n,\mathbb{R})$. For $A \in Gl(n,\mathbb{R})$, the left translation $M \longmapsto AM$ extends to a linear transformation of the whole linear space $Mat(n,\mathbb{R})$. The derivative of a linear transformation of a linear space is the linear transformation itself, if we identify the tangent spaces at different points with the linear space, so a left invariant vector field $X: Gl(n,\mathbb{R}) \longrightarrow Mat(n,\mathbb{R})$ has the form $X(A) = AM$, where $M \in Mat(n,\mathbb{R})$ is a fixed matrix.

The integral curves of a left invariant vector field on $Gl(n,\mathbb{R})$ can be described with the help of the exponential function for matrices. If M is an arbitrary square matrix, then we define e^M as the sum

$$\sum_{n=0}^{\infty} \frac{M^n}{n!} .$$

If we define the curve $\gamma_A: \mathbb{R} \longrightarrow Gl(n,\mathbb{R})$ by

$$\gamma_A(t) = A\, e^{Mt},$$

then we obtain an integral curve of the vector field $X(A) = AM$. Indeed,

$$\gamma_A'(t) = Ae^{Mt}M = X(Ae^{Mt}) = X(\gamma_A(t)).$$

The flow generated by the left invariant vector field X consists of the diffeomorphisms

$$H_t(A) = Ae^{Mt},$$

that is, H_t is a right translation by e^{Mt}.

Now let us take two left invariant vector fields $X(A)=AM$ and $Y(A)=AN$ and consider the flows H_t and G_t generated by them.

Computing $G_t \circ H_t(A) - H_t \circ G_t(A)$ up to $o(t^2)$, we get

$$G_t \circ H_t(A) - H_t \circ G_t(A) = A(e^{Mt}e^{Nt} - e^{Nt}e^{Mt}) = A(I+Mt+\frac{1}{2}(Mt)^2)(I+Nt+\frac{1}{2}(Nt)^2) - $$
$$-A(I+Nt+\frac{1}{2}(Nt)^2)(I+Mt+\frac{1}{2}(Mt)^2) + o(t^2) =$$
$$A(MN-NM)t^2 + o(t^2).$$

We obtain, that *the Lie algebra of* $Gl(n,\mathbb{R})$ *is isomorphic to the Lie algebra of all matrices with Lie bracket* $[M,N] = MN-NM$.

Further Exercises

Exercise 10-1. Let ∂_1 and ∂_2 be the two coordinate vector fields on $\mathbb{R}^2$ determined by the identity mapping. Describe the vector fields

$$X(x_1,x_2) = x_1\partial_1 + x_2\partial_2$$
$$Y(x_1,x_2) = x_2\partial_1 - x_1\partial_2,$$

compute their Lie bracket, and determine the flows generated by them.

Exercise 10-2. Show that the Lie algebra of $SO(n)$ is isomorphic to the Lie algebra of skew-symmetric $n \times n$ matrices with Lie bracket $[X,Y]=XY-YX$.

Exercise 10-3. Show that $\mathbb{R}^3$ endowed with the cross-product x is a 3-dimensional Lie algebra isomorphic to the Lie algebra of $SO(3)$.

Exercise 10-4. For $\mathbf{v} \in \mathbb{R}^3$, let $X_{\mathbf{v}}$ denote the vector field on $\mathbb{R}^3$, defined by

$$X_{\mathbf{v}}(\mathbf{x}) = \mathbf{v} \times \mathbf{x}.$$

Describe the flow generated by $X_{\mathbf{v}}$, and prove that

$$[X_{\mathbf{v}}, X_{\mathbf{w}}] = -X_{\mathbf{v}\times\mathbf{w}}.$$

Exercise 10-5. Show that the Lie algebra of left invariant vector fields on a Lie group is isomorphic to the Lie algebra of right invariant vector fields.

Unit 11. Differentiation of Vector Fields

Affine connection at a point, global affine connection, Christoffel symbols, covariant derivation of vector fields along a curve, parallel vector fields and parallel translation, symmetric connections, Riemannian manifolds, compatibility with a Riemannian metric, the fundamental theorem of Riemannian geometry, Levi-Civita connection.

Although there is a natural way to differentiate a smooth function defined on a manifold with respect to a tangent vector, there is no natural way to differentiate vector fields. In fact, there are lot of possible rules for differentiating vector fields with respect to a tangent vector, and to choose one of them, (the most appropriate one), the differentiable manifold structure alone is not enough. A fixed rule for the differentiation of vector fields is itself an additional structure on the manifold, called an affine connection. Later we shall see that on Riemannian manifolds i.e. on manifolds the tangent spaces of which are equipped with a dot product we can introduce differentiation of vector fields in a natural way. A precise formulation of this statement is the "fundamental theorem of Riemannian geometry".

As far as only vector fields on an open domain of $\mathbb{R}^n$ are considered, the following definition seems to be quite natural.

Definition. The derivative of a smooth vector field X on an open subset U in $\mathbb{R}^n$ with respect to a tangent vector $Y \in T_p\mathbb{R}^n$ is defined by

$$\nabla_Y X = (X\circ\gamma)'(0) ,$$

where $\gamma : [-\varepsilon.\varepsilon] \longrightarrow U$ is any smooth curve such that $\gamma(0) = p$ and $\gamma'(0) = Y$. We see that

$$\nabla_Y X = \sum_{i=1}^{n} Y(X_i)\, \partial_i(p) , \qquad (*)$$

where ∂_i denotes the i-th coordinate vector field on $\mathbb{R}^n$, X_i are the components of the vector field X. In particular, the value of $\nabla_Y X$ does not depend on the choice of γ.

It is easy to check that differentiation of vector fields has the following properties.

$$\nabla_{(Y_1+Y_2)}X = \nabla_{Y_1}X + \nabla_{Y_2}X \tag{1}$$

$$\nabla_{cY}X = c\,\nabla_Y X \tag{2}$$

$$\nabla_Y(X_1+X_2) = \nabla_Y X_1 + \nabla_Y X_2 \tag{3}$$

$$\nabla_Y(f\,X) = Y(f)\,X + f\,\nabla_Y X \tag{4}$$

$$\nabla_{X_1}X_2 - \nabla_{X_2}X_1 = [X_1,X_2] \quad \text{(torsion free property)} \tag{5}$$

$$Y(\langle X_1,X_2\rangle) = \langle\nabla_Y X_1,X_2\rangle+\langle X_1,\nabla_Y X_2\rangle \quad \text{(agreement with the metric)} \tag{6}$$

where $X_1,X_2 \in \mathfrak{X}(\mathbb{R}^n)$, $Y \in T_p\mathbb{R}^n$, $f \in \mathcal{F}(\mathbb{R}^n)$, $c \in \mathbb{R}$.

Now we shall study the general case. Let M be a smooth manifold. As we mentioned, there is no natural rule for derivation of vector fields on M, so we introduce such rules axiomatically, as operations satisfying some of the properties (1-6).

Definition. An affine connection at a point $p \in M$ is a mapping which assigns to each tangent vector $Y \in T_pM$ and each vector field $X \in \mathfrak{X}(M)$ a new tangent vector $\nabla_Y X \in T_pM$ called the covariant derivative of X with respect to Y and satisfies the following identities

$$\nabla_{(Y_1+Y_2)}X = \nabla_{Y_1}X + \nabla_{Y_2}X \tag{1}$$

$$\nabla_{cY}X = c\,\nabla_Y X \tag{2}$$

$$\nabla_Y(X_1+X_2) = \nabla_Y X_1 + \nabla_Y X_2 \tag{3}$$

$$\nabla_Y(f\,X) = Y(f)\,X + f(p)\,\nabla_Y X \tag{4}$$

where $X_1,X_2 \in \mathfrak{X}(M)$, $Y,Y_1,Y_2 \in T_pM$, $f \in \mathcal{F}(M)$, $c \in \mathbb{R}$.

Definition. A global affine connection (or briefly a connection) on M is a mapping which assigns to two smooth vector fields Y and X a new one $\nabla_Y X$ called the covariant derivative of the vector field X with respect to the vector field Y, having the following properties

$$\nabla_{(Y_1+Y_2)}X = \nabla_{Y_1}X + \nabla_{Y_2}X \tag{1'}$$

$$\nabla_{cY}X = c\,\nabla_Y X \tag{2'}$$

$$\nabla_Y(X_1+X_2) = \nabla_Y X_1 + \nabla_Y X_2 \tag{3'}$$

$$\nabla_Y(f\,X) = Y(f)\,X + f\,\nabla_Y X \tag{4'}$$

where $X_1, X_2, Y, Y_1, Y_2 \in \mathfrak{X}(M)$, $f, c \in \mathcal{F}(M)$.

Lemma. For a global affine connection ∇, $X, Y \in \mathfrak{X}(M)$, $p \in M$, the tangent vector $(\nabla_Y X)(p)$ depends only on the behavior of X and Y in an open neighborhood of p.

Proof. Let us suppose that the vector fields X_1 and X_2 coincide on an open neighborhood U of p. Choose a smooth function $h \in \mathcal{F}(M)$ which is zero outside U and constant 1 on a neighborhood of p. Then we have $h(X_1 - X_2) = 0$, consequently

$$0 = \nabla_Y(h(X_1-X_2)) = h\nabla_Y(X_1-X_2)+Y(h)(X_1-X_2).$$

Computing the right hand side at p we get

$$0 = (\nabla_Y X_1)(p) - (\nabla_Y X_2)(p).$$

Similarly, if the vector fields Y_1 and Y_2 coincide on an open neighborhood U of p and h is chosen as above, then we have

$$0 = h(Y_1-Y_2)X = \nabla_{h(Y_1-Y_2)}X = h\nabla_{Y_1} X - h\nabla_{Y_2} X \ ,$$

which yields $(\nabla_{Y_1} X)(p) - (\nabla_{Y_2} X)(p)$. ∎

The lemma shows that an affine connection can be restricted onto any open subset and can be recovered from its restrictions onto the elements of an open cover.

Let $x_1, \dots, x_n$ be local coordinates defined on an open subset U of M and $\partial_1 = \frac{\partial}{\partial x_1}, \quad , \partial_n = \frac{\partial}{\partial x_n}$ be the corresponding basis vector fields on U. Given an affine connection ∇ on M, we can express the vector field $\nabla_{\partial_i} \partial_j$ as a linear combination of the basis vector fields

$$\nabla_{\partial_i} \partial_j = \sum_{k=1}^{n} \Gamma^k_{ji} \, \partial_k \ .$$

The components Γ^k_{ji} are smooth functions called Christoffel symbols.

Proposition. The restriction of an affine connection onto an open coordinate neighborhood U is uniquely determined by the Christoffel symbols. Any n^3 smooth functions Γ^k_{ji} on U may be Christoffel symbols for an appropriate affine connection on U.

Proof. Let $X = \sum_{i=1}^{n} X_i \, \partial_i$, $Y = \sum_{j=1}^{n} Y_j \, \partial_j$ be two smooth vector fields on U. Then by the properties of affine connections $\nabla_Y X$ can be computed as follows

$$\nabla_Y X = \nabla_{\left(\sum\limits_{j=1}^{n} Y_j \partial_j\right)} \left(\sum_{i=1}^{n} X_i \partial_i\right) = \sum_{j=1}^{n} Y_j \nabla_{\partial_j} \left(\sum_{i=1}^{n} X_i \partial_i\right) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} Y_j \left(\partial_j(X_i)\partial_i + X_i \nabla_{\partial_j} \partial_i\right) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \left(Y_j \partial_j(X_i)\partial_i + X_i Y_j \sum_{k=1}^{n} \Gamma_{ij}^{k} \partial_k\right) =$$

$$= \sum_{k=1}^{n} \left(Y(X_k) + \sum_{i=1}^{n} \sum_{j=1}^{n} X_i Y_j \Gamma_{ij}^{k}\right) \partial_k .$$

This formula shows that the knowledge of the Christoffel symbols enables us to compute the covariant derivative of any vector field with respect to any other one. On the other hand, if Γ_{ij}^{k} is an arbitrary smooth function on U for $1 \leqslant i,j,k \leqslant n$, then defining the covariant derivative of a vector field by the above formula, we obtain an affine connection on U. $\blacksquare$

Observe, that in fact, the tangent vector $(\nabla_Y X)(p)$ depends only on the vector $Y(p)$, so a global affine connection on a manifold defines an affine connection at each of its points. Furthermore, we do not need to know the vector field X everywhere on U to compute $(\nabla_Y X)(p)$. It is enough to know X at the points of a curve γ $[-\varepsilon,\varepsilon] \longrightarrow M$ such that $\gamma(0) = p$, $\gamma'(0)=Y(p)$.

<u>Definition</u>. Let $\gamma: [a,b] \longrightarrow M$ be a smooth curve in M. A smooth <u>vector field</u> X <u>along the curve</u> γ is a smooth mapping $X: [a,b] \longrightarrow TM$ which assigns to each $t \in [a,b]$ a tangent vector $X(t) \in T_{\gamma(t)}M$.

Now suppose that M is provided with an affine connection. Then any vector field X along γ determines a new vector field $\frac{DX}{dt}$ along γ called the <u>covariant derivative</u> of X. In terms of local coordinates, if $\partial_1, \dots, \partial_n$ are the basis vector fields determined by a chart and the functions $X_i: [a,b] \longrightarrow \mathbb{R}$ and $Y_i: [a,b] \longrightarrow \mathbb{R}$ assign to $t \in [a,b]$ the components of the vectors $X(\gamma(t))$ and $D_{\gamma'(t)}$ in the basis $\partial_1(\gamma(t)), \dots, \partial_n(\gamma(p))$, then $\frac{DX}{dt}$ is defined as follows

$$\frac{DX}{dt}(\tau) := \sum_{k=1}^{n} \left(X_k'(\tau) + \sum_{i=1}^{n} \sum_{j=1}^{n} X_i(\tau) Y_j(\tau) \Gamma_{ij}^{k}(\gamma(\tau))\right) \partial_k(\gamma(\tau)).$$

<u>Definition</u>. A vector field X along a curve γ is said to be a <u>parallel vector field</u> if the covariant derivative $\frac{DX}{dt}$ is identically zero.

Proposition. Given a curve γ and a tangent vector X_0 at the point $\gamma(0)$, there is a unique parallel vector field X along γ which extends X_0.

Proof. The proposition follows from results on ordinary differential equations. Using local coordinates, condition $\frac{DX}{dt} = 0$ yields a system of ordinary differential equations for the components of X

$$X_k' + \sum_{i=1}^{n} \sum_{j=1}^{n} X_i Y_j \Gamma^k_{ij} \circ \gamma = 0.$$

Since these equations are linear, the existence and uniqueness theorem for linear differential equations guaranties that the solutions of this system of differential equations are uniquely determined by the initial values $X_k(0)$ and can be defined for all relevant values of t. ▮

The vector $X_t = X(\gamma(t))$ is said to be obtained from X_0 by parallel translation along γ.

Definition. A connection is called symmetric or torsion free if it satisfies the identity

$$\nabla_X Y - \nabla_Y X = [X,Y].$$

Applying this identity to the case $X = \partial_i$, $Y = \partial_j$, since $[\partial_i, \partial_j] = 0$ one obtains the relation

$$\Gamma^k_{ij} = \Gamma^k_{ji}.$$

Conversely, if $\Gamma^k_{ij} = \Gamma^k_{ji}$ then using the expression of covariant derivative with the help of Christoffel symbols we get

$$\nabla_X Y - \nabla_Y X =$$

$$= \sum_{k=1}^{n} \left(X(Y_k) + \sum_{i=1}^{n} \sum_{j=1}^{n} Y_i X_j \Gamma^k_{ij} \right) \partial_k - \sum_{k=1}^{n} \left(Y(X_k) + \sum_{i=1}^{n} \sum_{j=1}^{n} X_i Y_j \Gamma^k_{ij} \right) \partial_k$$

$$= \sum_{k=1}^{n} \left(X(Y_k) - Y(X_k) \right) \partial_k$$

$$= [X,Y].$$

There is a useful characterization of symmetry. Consider a "parameterized surface" in M that is a smooth mapping $s: R \longrightarrow M$ from a rectangular domain R of the plane $\mathbb{R}^2$ into M.

By a vector field X along s is meant a mapping which assigns to each $(x,y) \in R$ a tangent vector $X(x,y) \in T_{s(x,y)}M$.

As examples, the two standard vector fields $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ on the plane give rise to vector fields $Ts(\frac{\partial}{\partial x})$ and $Ts(\frac{\partial}{\partial y})$ along s. These will be denoted briefly by $\frac{\partial s}{\partial x}$ and $\frac{\partial s}{\partial y}$. Here $Ts: T\mathbb{R}^2 \longrightarrow TM$ denotes the derivative of s.

For any smooth vector field X along s the <u>covariant partial derivatives</u> $\frac{DX}{\partial x}$, $\frac{DX}{\partial y}$ are new vector fields along s constructed as follows. For each fixed y_0, restricting X to the curve $x \longmapsto s(x,y_0)$ one obtains a vector field along this curve. Its covariant derivative with respect to x is defined to be $\frac{DX}{\partial x}(x,y_0)$. This defines $\frac{DX}{\partial x}$ along the entire parameterized surface s.

<u>Proposition</u>. A connection is symmetric if and only if

$$\frac{D}{\partial x}\frac{\partial s}{\partial y} = \frac{D}{\partial y}\frac{\partial s}{\partial x}$$

for any parameterized surface s in M.

<u>Proof</u>. Let us choose a local coordinate system $(x_1,...,x_n)$ on M. The mapping s is given by n functions $s_i = x_i \circ s$. The vector field $\frac{\partial s}{\partial y}$ has the form

$$\frac{\partial s}{\partial y} = \sum_{i=1}^{n} \frac{\partial s_i}{\partial y}(\partial_i \circ s).$$

The partial covariant derivative of this vector field with respect to x is equal to

$$\frac{D}{\partial x}\frac{\partial s}{\partial y} = \frac{D}{\partial x}\left(\sum_{i=1}^{n}\frac{\partial s_i}{\partial y}(\partial_i \circ s)\right) = \sum_{i=1}^{n}\left(\frac{\partial^2 s_i}{\partial x \partial y}\partial_i \circ s + \frac{\partial s_i}{\partial y}\frac{D}{\partial x}(\partial_i \circ s)\right) =$$

$$= \sum_{i=1}^{n}\frac{\partial^2 s_i}{\partial x \partial y}\partial_i \circ s + \sum_{k=1}^{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial s_i}{\partial y}\frac{\partial s_j}{\partial x}\Gamma_{ij}^{k} \circ s\right)\partial_k \circ s.$$

This formula shows that interchanging the role of x and y we obtain the same vector field for any s if and only if $\Gamma_{ij}^{k} = \Gamma_{ji}^{k}$. ∎

Roughly speaking, the torsion free condition halves the degree of freedom in the choice of Christoffel symbols, a symmetric connection is uniquely determined by $n\frac{n(n+1)}{2}$ arbitrarily chosen functions, nevertheless, the space of symmetric affine connections on a manifold is still infinite dimensional.

We can reduce further the degree of freedom putting condition (6) on the connection. This condition however does not make sense on an arbitrary manifold, because dot product of tangent vectors at a given point is not defined in general.

<u>Definition</u>. Let M be a differentiable manifold. Suppose that each tangent space T_pM of M is equipped with a positive definite symmetric bilinear form $\langle , \rangle_p$ so that for any two smooth vector fields $X,Y: M \longrightarrow TM$ the function $M \ni p \longmapsto \langle X(p),Y(p)\rangle_p$ is smooth. Then M together with this structure is a

Riemannian manifold, the system of bilinear forms on the tangent spaces is the Riemannian metric on M.

Example. A hypersurface of $\mathbb{R}^n$ together with the first fundamental form is a Riemannian manifold.

Definition. A connection ∇ on M is compatible with the Riemannian metric if parallel translation preserves inner products. In other words, for any curve γ and any pair X, Y of parallel vector fields along γ, the inner product $\langle X, Y\rangle$ should be constant.

Lemma. Suppose that the connection is compatible with the metric. Let V, W be any two vector fields along γ. Then

$$\frac{d}{dt}\langle V,W\rangle = \left\langle \frac{DV}{dt}, W\right\rangle + \left\langle V, \frac{DW}{dt}\right\rangle .$$

Proof. Choose parallel vector fields $X_1,\ldots,X_n$ along γ which are orthonormal at one point of γ and hence at every point of γ. Then the given fields V and W can be expressed as $\sum_{i=1}^{n} v_i X_i$ and $\sum_{i=1}^{n} w_i X_i$ respectively (where $v_i = \langle V, X_i\rangle$ is a real valued function). It follows that $\langle V, W\rangle = \sum_{i=1}^{n} v_i w_i$ and that

$$\frac{DV}{dt} = \sum_{i=1}^{n} \frac{dv_i}{dt} X_i, \qquad \frac{DW}{dt} = \sum_{i=1}^{n} \frac{dw_i}{dt} X_i .$$

Therefore

$$\left\langle \frac{DV}{dt}, W\right\rangle + \left\langle V, \frac{DW}{dt}\right\rangle = \sum_{i=1}^{n}\left(\frac{dv_i}{dt} w_i + v_i \frac{dw_i}{dt}\right) = \frac{d}{dt}\langle V,W\rangle. \blacksquare$$

Corollary. An affine connection on a Riemannian manifold is compatible with the metric if and only if for any vector fields X_1, X_2 on M and any tangent vector $Y \in T_pM$ we have

$$Y(\langle X_1,X_2\rangle) = \langle \nabla_Y X_1, X_2\rangle + \langle X_1, \nabla_Y X_2\rangle.$$

Theorem. (Fundamental theorem of Riemannian geometry.) A Riemannian manifold possesses one and only one symmetric connection which is compatible with its metric.

Proof. Applying the compatibility condition to the basis vector fields $\partial_i, \partial_j, \partial_k$, corresponding to a fixed chart on the manifold and setting $\langle\partial_j, \partial_k\rangle = g_{jk}$ one obtains the identity

$$\partial_i \, g_{jk} = \langle \nabla_{\partial_i}\partial_j \, , \, \partial_k \rangle + \langle \partial_j \, , \, \nabla_{\partial_i}\partial_k \rangle.$$

permuting i, j and k this gives three linear equations relating the three quantities

$$\langle \nabla_{\partial_i}\partial_j \, , \, \partial_k \rangle , \quad \langle \nabla_{\partial_j}\partial_k \, , \, \partial_i \rangle , \quad \langle \nabla_{\partial_k}\partial_i \, , \, \partial_j \rangle.$$

(There are only three such quantities since $\nabla_{\partial_i}\partial_j = \nabla_{\partial_j}\partial_i$.) These equations can be solved uniquely; yielding the first Christoffel identity

$$\langle \nabla_{\partial_i}\partial_j \, , \, \partial_k \rangle = \frac{1}{2} \left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right) .$$

The left hand side of this identity is equal to $\sum_{l=1}^{n} \Gamma^l_{ij} \, g_{lk}$. Multiplying by the inverse (g^{kl}) of the matrix (g_{lk}) this yields the second Christoffel identity

$$\Gamma^l_{ij} = \sum_{k=1}^{n} \frac{1}{2} \left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right) g^{kl}.$$

Thus the connection is uniquely determined by the metric.

Conversely, defining Γ^l_{ij} by this formula, one can verify that the resulting connection is symmetric and compatible with the metric. This completes the proof. ▮

The unique symmetric affine connection which is compatible with the metric on a Riemannian manifold is called the Levi-Civita connection.

The connection ∇ we introduced on open subsets of $\mathbb{R}^n$ is just the Levi-Civita connection of $\mathbb{R}^n$.

Consider now a parameterized hypersurface $\mathbf{r}:\Omega \longrightarrow \mathbb{R}^n$. It is a Riemannian manifold with respect to the first fundamental form. The basis vector fields $\mathbf{r}_i$ through suitable identifications are the same as the basis vector fields ∂_i corresponding to the chart $\mathbf{r}^{-1}$. Comparing the formulae

$$\partial_{\mathbf{r}_i}\mathbf{r}_j = \mathbf{r}_{ij} = \sum_{k=1}^{n-1} \Gamma^k_{ij} \, \mathbf{r}_k + h_{ij} \, \mathbf{N}$$

and

$$\Gamma^k_{ij} = \frac{1}{2} \sum_{l=1}^{n-1} g^{kl}\left(g_{li,j} + g_{lj.i} - g_{ij,l}\right)$$

proved in unit 7 with the formulae derived for the Levi-Civita connection we may conclude that the Christoffel symbols of a hypersurface introduced previously in unit 7 are the Christoffel symbols of the Levi-Civita connection of the hypersurface. Furthermore, denoting by $\tilde{\nabla}$ the Levi-Civita connection of the hypersurface, we see that for tangential vector fields X,Y $\tilde{\nabla}_Y X$ is the tangential component of $\partial_Y X$.

$$\tilde{\nabla}_Y X = \partial_Y X - \langle \partial_Y X , \mathbf{N} \rangle \mathbf{N}$$

Further Exercises

Exercise 11-1. Show that if ∇ and $\tilde{\nabla}$ are two affine connections on a manifold M, then their difference $S(X,Y) = \nabla_X Y - \tilde{\nabla}_X Y$ is an $\mathcal{F}(M)$-bilinear mapping. (In other words, S is a tensor field of valency (1,2)). Conversely, the sum of a connection and an $\mathcal{F}(M)$-bilinear mapping $S:\mathfrak{X}(M)\times\mathfrak{X}(M)\longrightarrow\mathfrak{X}(M)$ is a connection. (According to these statements, affine global connections form an affine space over the linear space of (1,2)-tensor fields.)

Exercise 11-2. Show that the <u>torsion</u> $T:\mathfrak{X}(M)\times\mathfrak{X}(M)\longrightarrow\mathfrak{X}(M)$ of a connection ∇ defined by

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$$

is an $\mathcal{F}(M)$-bilinear mapping (i.e., T is a tensor field).

Exercise 11-3. Show that if T is the torsion of an affine connection ∇ the $\nabla - T/2$ is a symmetric connection.

Exercise 11-4. Check that the connection defined by the Christoffel symbols

$$\Gamma^k_{ij} = \frac{1}{2} \sum_{l=1}^{n-1} g^{kl}(g_{li,j} + g_{lj.i} - g_{ij,l})$$

is symmetric and compatible with the metric.

Unit 12. Curvature

Curvature operator, curvature tensor, Bianchi identities, Riemann-Christoffel tensor, symmetry properties of the Riemann-Christoffel tensor, sectional curvature, Schur's theorem, space forms, Ricci tensor, Ricci curvature, scalar curvature, curvature tensor of a hypersurface.

If ∇ is an affine connection on a manifold M, then we may consider the operator

$$R(X,Y) = [\nabla_X,\nabla_Y] - \nabla_{[X,Y]}:\mathfrak{X}(M)\longrightarrow \mathfrak{X}(M),$$

where $[\nabla_X,\nabla_Y]=\nabla_X\circ\nabla_Y-\nabla_Y\circ\nabla_X$ is the usual commutator of operators. The mapping that assigns to the vector fields X,Y the operator R(X,Y) is called the <u>curvature operator</u> of the connection. The assignment

$$\mathfrak{X}(M)\times\mathfrak{X}(M)\times\mathfrak{X}(M)\longrightarrow \mathfrak{X}(M)$$
$$(X,Y,Z)\longmapsto R(X,Y)(Z)$$

is called the <u>curvature tensor</u> of the connection. To reduce the number of brackets, we shall denote R(X,Y)(Z) simply by R(X,Y;Z). Thus, the letter R is used in two different meanings, later it will denote also a third mapping, but the number of arguments of R makes always clear which meaning is considered.

<u>Proposition</u>. The curvature tensor is linear over the ring of smooth functions in each of its arguments, and it is skew symmetric in the first two arguments.

<u>Proof</u>. Skew symmetry in the first two arguments is clear, since

$$R(X,Y) = [\nabla_X,\nabla_Y] - \nabla_{[X,Y]} = -[\nabla_Y,\nabla_X] + \nabla_{[Y,X]} = -R(Y,X).$$

According to this, it suffices to check linearity of the curvature tensor in the first and third arguments.

Linearity in the first argument is proved by the following identities.

$$R(X_1+X_2,Y) = [\nabla_{X_1+X_2},\nabla_Y] - \nabla_{[X_1+X_2,Y]} = [\nabla_{X_1}+\nabla_{X_2},\nabla_Y] - \nabla_{[X_1,Y]+[X_2,Y]} =$$
$$= [\nabla_{X_1},\nabla_Y] + [\nabla_{X_2},\nabla_Y] - \nabla_{[X_1,Y]}- \nabla_{[X_2,Y]} = R(X_1,Y) + R(X_2,Y).$$

and

$$R(fX,Y;Z) = ([\nabla_{fX},\nabla_Y] - \nabla_{[fX,Y]})(Z) = f\nabla_X\nabla_YZ-\nabla_Y(f\nabla_XZ)-\nabla_{f[X,Y]-Y(f)X}(Z) =$$
$$= f\nabla_X\nabla_YZ - f\nabla_Y\nabla_XZ - Y(f)\nabla_XZ -f\nabla_{[X,Y]}Z + Y(f)\nabla_X(Z) =$$
$$= f(\nabla_X\nabla_YZ - \nabla_Y\nabla_XZ - \nabla_{[X,Y]}Z) = f\, R(X,Y;Z).$$

Additivity in the third argument is clear, since R(X,Y) is built up of the additive operators ∇_X, ∇_Y and their compositions. To have linearity, we need

$$R(X,Y;fZ) = \nabla_X\nabla_Y(fZ) - \nabla_Y\nabla_X(fZ) - \nabla_{[X,Y]}(fZ) =$$
$$= \nabla_X(Y(f)Z+f\nabla_YZ)-\nabla_Y(X(f)Z+f\nabla_XZ)-[X,Y](f)Z-f\nabla_{[X,Y]}Z =$$
$$= XY(f)Z+Y(f)\nabla_XZ+X(f)\nabla_YZ+f\nabla_X\nabla_YZ -YX(f)Z-X(f)\nabla_YZ-Y(f)\nabla_XZ-f\nabla_Y\nabla_XZ-$$
$$-XY(f)Z+YX(f)Z - f\nabla_{[X,Y]}Z$$
$$= f(\nabla_X\nabla_YZ - \nabla_Y\nabla_XZ - \nabla_{[X,Y]}Z) = f\ R(X,Y;Z). \blacksquare$$

The proposition is a bit surprising, because the curvature tensor is built up from covariant derivations, which are not linear operators over the ring of smooth functions.

We have already introduced tensor fields over a hypersurface. We can introduce tensor fields over a manifold in the same manner. A tensor field T of type (k,ℓ) is an assignment to every point p of a manifold M a tensor $T(p)$ of type (k,ℓ) over the tangent space T_pM. If $\partial_1,\dots,\partial_n$ are the basis vector fields defined by a chart over the domain of the chart, and we denote by $dx^1(p)$, ..., $dx^n(p)$ the dual basis of $\partial_1(p),\dots,\partial_n(p)$, then a tensor field is uniquely determined over the domain of the chart by the components

$$T^{i_1\dots i_k}_{j_1\dots j_l}(p)=T(p)(dx^{i_1}(p),\dots,dx^{i_k}(p);\partial_{j_1}(p),\dots,\partial_{j_l}(p)).$$

We say that the tensor field is <u>smooth</u>, if for any chart the functions $T^{i_1\dots i_k}_{j_1\dots j_l}$ are smooth. We shall consider only smooth tensor fields.

Tensors of valency (1,0) are the vector fields, tensors of valency (0,1) are the <u>differential 1-forms</u>. Thus, a differential 1-form assigns to every point of the manifold a linear function on the tangent space at that point. Differential 1-forms form a module over the ring of smooth functions, which we denote by $\Omega^1(M)$.

Every tensor field defines a multi-$\mathcal{F}(M)$-linear mapping

$$\Omega^1(M)\times\dots\times\Omega^1(M)\times\mathfrak{X}(M)\times\dots\times\mathfrak{X}(M) \longrightarrow \mathcal{F}(M)$$

and conversely, every such multi-$\mathcal{F}(M)$-linear mapping comes from a tensor field. (Check this!) Therefore, tensor fields can be identified with multi-$\mathcal{F}(M)$-linear mappings $\Omega^1(M)\times\dots\times\Omega^1(M)\times\mathfrak{X}(M)\times\dots\times\mathfrak{X}(M) \longrightarrow \mathcal{F}(M)$.

Tensors of type (1,k), that is multi-$\mathcal{F}(M)$-linear mappings

$$\Omega^1(M)\times\mathfrak{X}(M)\times\dots\times\mathfrak{X}(M) \longrightarrow \mathcal{F}(M)$$

can be identified in a natural way with multi-$\mathcal{F}(M)$-linear mappings

$$\mathfrak{X}(M)\times\dots\times\mathfrak{X}(M) \longrightarrow \mathfrak{X}(M).$$

By this identification, $R:\mathfrak{X}(M)\times\ldots\times\mathfrak{X}(M) \longrightarrow \mathfrak{X}(M)$ corresponds to $\tilde{R}:\Omega^1(M)\times\mathfrak{X}(M)\times\ldots\times\mathfrak{X}(M) \longrightarrow \mathcal{F}(M)$, defined by $\tilde{R}(\omega;X_1,\ldots,X_k)=\omega(R(X_1,\ldots,X_k))$.

Using these identifications, the curvature tensor is a tensor field of valency (1,3) by the proposition. It is a remarkable consequence, that although the vectors $\nabla_X Z(p)$ and $\nabla_Y Z(p)$ are not determined by the vectors $X(p),Y(p),Z(p)$, to compute the value of $R(X,Y;Z)$ at p it suffices to know $X(p),Y(p),Z(p)$.

Beside skew-symmetry in the first two arguments, the curvature tensor has many other symmetry properties.

Theorem. (First Bianchi Identity). If R is the curvature tensor of a torsion free connection, then

$$R(X,Y;Z) + R(Y,Z;X) + R(Z,X;Y) = 0$$

for any three vector fields X,Y,Z.

Proof. Let us introduce the following notation. If $F(X,Y,Z)$ is a function of the vector fields X,Y,Z, then denote by $\circlearrowleft F(X,Y,Z)$ or $\underset{XYZ}{\circlearrowleft} F(X,Y,Z)$ the sum of the values of F at all cyclic permutations of the variables (X,Y,Z)

$$\circlearrowleft F(X,Y,Z) = F(X,Y,Z) + F(Y,Z,X) + F(Z,X,Y).$$

We shall use several times that behind the cyclic summation $\circlearrowleft$ we may cyclically rotate X,Y,Z in any expression

$$\circlearrowleft F(X,Y,Z) = \circlearrowleft F(Y,Z,X) = \circlearrowleft F(Z,X,Y).$$

The theorem claims vanishing of

$$\circlearrowleft R(X,Y;Z) = \circlearrowleft (\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z) = \circlearrowleft (\nabla_X\nabla_Y Z - \nabla_X\nabla_Z Y - \nabla_{[X,Y]}Z)$$
$$= \circlearrowleft (\nabla_X [Y,Z] - \nabla_{[X,Y]}Z) = \circlearrowleft (\nabla_Z [X,Y] - \nabla_{[X,Y]}Z) = \circlearrowleft [Z,[X,Y]],$$

but the latter expression is 0 according to the Jacobi identity on the Lie bracket of vector fields. (At the third and fifth equality we used the torsion free property of ∇.) ∎

The presence of an affine connection on a manifold allows us to differentiate not only vector fields, but also tensor fields of any type.

Definition. Let (M,∇) be a manifold with an affine connection. If $\omega \in \Omega^1(M)$ is a 1-form, X is a vector field, then we define the *covariant derivative* $\nabla_X\omega$ *of* ω with respect to X to be the 1-form

$$(\nabla_X\omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y), \quad Y\in\mathfrak{X}(M).$$

In general, the *covariant derivative* $\nabla_X T$ *of a tensor field*

$$T:\Omega^1(M)\times\ldots\times\Omega^1(M)\times\mathfrak{X}(M)\times\ldots\times\mathfrak{X}(M) \longrightarrow \mathcal{F}(M)$$

of valency (k,ℓ) with respect to a vector field X is a tensor field of the same valency, defined by

$$(\nabla_X T)(\omega_1,\dots,\omega_k;X_1,\dots,X_\ell) = X(T(\omega_1,\dots,\omega_k;X_1,\dots,X_\ell)) - \\ - \sum_{i=1}^{k} T(\omega_1,\dots,\nabla_X\omega_i,\dots,\omega_k;X_1,\dots,X_\ell) - \\ - \sum_{j=1}^{\ell} T(\omega_1,\dots,\omega_k;X_1,\dots,\nabla_X X_j,\dots,X_\ell).$$

For the case of the curvature tensor, this definition gives

$$(\nabla_X R)(Y,Z;W) = \nabla_X(R(Y,Z;W)-R(\nabla_X Y,Z;W)-R(Y,\nabla_X Z;W)-R(Y,Z;\nabla_X W).$$

<u>Theorem</u>. (Second Bianchi Identity) The curvature tensor of a torsion free connection satisfies the following identity

$$\underset{XYZ}{\circlearrowleft} (\nabla_X R)(Y,Z;W) = (\nabla_X R)(Y,Z;W) + (\nabla_Y R)(Z,X;W) + (\nabla_Z R)(X,Y;W) = 0.$$

<u>Proof</u>. $(\nabla_X R)(Y,Z;W)$ is the value of the operator

$$\nabla_X\circ R(Y,Z)-R(\nabla_X Y,Z)-R(Y,\nabla_X Z)-R(Y,Z)\circ\nabla_X : \mathfrak{X}(M) \longrightarrow \mathfrak{X}(M)$$

on the vector field W, hence we have to prove vanishing of the operator

$$\underset{XYZ}{\circlearrowleft} \nabla_X\circ R(Y,Z)-R(\nabla_X Y,Z)-R(Y,\nabla_X Z)-R(Y,Z)\circ\nabla_X .$$

First, we have

$$\underset{XYZ}{\circlearrowleft}\nabla_X\circ R(Y,Z)-R(Y,Z)\circ\nabla_X=\underset{XYZ}{\circlearrowleft}(\nabla_X\nabla_Y\nabla_Z-\nabla_X\nabla_Z\nabla_Y-\nabla_X\nabla_{[Y,Z]})-(\nabla_Y\nabla_Z\nabla_X-\nabla_Z\nabla_Y\nabla_X-\nabla_{[Y,Z]}\nabla_X)=$$

$$\underset{XYZ}{\circlearrowleft}(\nabla_X\nabla_Y\nabla_Z-\nabla_X\nabla_Z\nabla_Y-\nabla_X\nabla_{[Y,Z]})-(\nabla_X\nabla_Y\nabla_Z-\nabla_X\nabla_Z\nabla_Y-\nabla_{[Y,Z]}\nabla_X)=\underset{XYZ}{\circlearrowleft} \nabla_{[Y,Z]}\nabla_X - \nabla_X\nabla_{[Y,Z]}.$$

On the other hand,

$$\underset{XYZ}{\circlearrowleft} -R(\nabla_X Y,Z)-R(Y,\nabla_X Z) = \underset{XYZ}{\circlearrowleft} R(\nabla_X Z,Y) - R(\nabla_X Y,Z) = \underset{XYZ}{\circlearrowleft} R(\nabla_Y X,Z) - R(\nabla_X Y,Z) =$$

$$\underset{XYZ}{\circlearrowleft} R(\nabla_Y X-\nabla_X Y,Z) = \underset{XYZ}{\circlearrowleft} R([Y,X],Z).$$

Combining these results,

$$\underset{XYZ}{\circlearrowleft} \nabla_X\circ R(Y,Z)-R(\nabla_X Y,Z)-R(Y,\nabla_X Z)-R(Y,Z)\circ\nabla_X$$

$$= \underset{XYZ}{\circlearrowleft} \nabla_{[Y,Z]}\nabla_X-\nabla_X\nabla_{[Y,Z]}+R([Y,X],Z)$$

$$= \underset{XYZ}{\circlearrowleft} \nabla_{[Y,Z]}\nabla_X-\nabla_X\nabla_{[Y,Z]}+R([Z,Y],X)$$

$$= \underset{XYZ}{\circlearrowleft} \nabla_{[Y,Z]}\nabla_X-\nabla_X\nabla_{[Y,Z]}+\nabla_{[Z,Y]}\nabla_X - \nabla_X\nabla_{[Z,Y]} - \nabla_{[[Z,Y],X]}$$

$$= \underset{XYZ}{\circlearrowleft} \nabla_{[[Y,Z],X]} = \nabla_{\circlearrowleft [[X,Y],Z]} = 0. \quad \blacksquare$$

In the remaining part of this unit, we shall deal with Riemannian manifolds. If $(M,\langle,\rangle)$ is a Riemannian, manifold with Levi-Civita connection ∇, and R is the curvature tensor of ∇, then we can introduce a tensor $\tilde{R}$ of valency $(0,4)$, related to R by the equation

$$\tilde{R}(X,Y;Z,W) = \langle R(X,Y;Z),W \rangle.$$

$\tilde{R}$ is the Riemann-Christoffel curvature tensor of the Riemannian manifold. To simplify notation, we shall denote $\tilde{R}$ also by R. This will not lead to confusion, since the Riemann-Christoffel tensor and the ordinary curvature tensor have different number of arguments.

Levi-Civita connections are connections of special type, so it is not surprising, that the curvature tensor of a Riemannian manifold has stronger symmetries than that of an arbitrary connection. Of course, the general results can be applied to Riemannian manifolds as well, and yield

$$R(X,Y;Z,W) = -R(Y,X;Z,W);$$

$$\underset{XYZ}{\circlearrowleft} R(X,Y;Z,W) = 0.$$

In addition to these symmetries, we have the following ones.

Theorem. The Riemann-Christoffel curvature tensor is skew-symmetric in the last two arguments

$$R(X,Y;Z,W) = - R(X,Y;W,Z).$$

Proof. By the compatibility of the connection and the metric, we have

$X(Y(\langle Z,W\rangle)) = X(\langle\nabla_Y Z,W\rangle+\langle Z,\nabla_Y W\rangle) = \langle\nabla_X\nabla_Y Z,W\rangle+\langle\nabla_Y Z,\nabla_X W\rangle+\langle\nabla_X Z,\nabla_Y W\rangle+\langle Z,\nabla_X\nabla_Y W\rangle,$

and similarly,

$Y(X(\langle Z,W\rangle)) = \langle\nabla_Y\nabla_X Z,W\rangle+\langle\nabla_X Z,\nabla_Y W\rangle+\langle\nabla_Y Z,\nabla_X W\rangle+\langle Z,\nabla_Y\nabla_X W\rangle.$

We also have

$[X,Y](\langle Z,W\rangle) = \langle\nabla_{[X,Y]}Z,W \rangle + \langle Z,\nabla_{[X,Y]}W \rangle.$

Subtracting from the first equality the second and the third one and applying $[X,Y] = X\circ Y - Y\circ X$, we obtain

$0 = \langle\nabla_X\nabla_Y Z-\nabla_Y\nabla_X Z-\nabla_{[X,Y]}Z,W\rangle+\langle Z,\nabla_X\nabla_Y W-\nabla_Y\nabla_X W-\nabla_{[X,Y]}W\rangle= R(X,Y;Z,W)+R(X,Y;W,Z).$ ∎

The symmetries we have by now imply a further symmetry.

Theorem. Assume that R is an arbitrary tensor of valency $(0,4)$ having the following symmetry properties

$$R(X,Y;Z,W) = -R(Y,X;Z,W) = -R(X,Y;W,Z)$$

and

$$\underset{XYZ}{\circlearrowleft} R(X,Y;Z,W) = 0.$$

Then

$$R(X,Y;Z,W) = R(Z,W;X,Y).$$

Proof. Let us apply the Bianchi identity for the following five arrangements

$$R(X,Y;Z,W) + R(Y,Z;X,W) + R(Z,X;Y,W) = 0$$
$$R(X,Y;W,Z) + R(Y,W;X,Z) + R(W,X;Y,Z) = 0$$
$$R(X,W;Z,Y) + R(W,Z;X,Y) + R(Z,X;W,Y) = 0$$
$$R(Y,Z;W,X) + R(Z,W;Y,X) + R(W,Y;Z,X) = 0$$
$$2R(Y,W;Z,X) + 2R(W,Z;Y,X) + 2R(Z,Y;W,X) = 0.$$

Changing the order of letters in the first two and last two places to the alphabetical order, we obtain the following equalities.

$$-R(X,Y;W,Z) - R(Y,Z;W,X) + R(X,Z;W,Y) = 0$$
$$R(X,Y;W,Z) - R(W,Y;X,Z) + R(W,X;Y,Z) = 0$$
$$R(W,X;Y,Z) + R(W,Z;X,Y) - R(X,Z;W,Y) = 0$$
$$R(Y,Z;W,X) + R(W,Z;X,Y) - R(W,Y;X,Z) = 0$$
$$2R(W,Y;X,Z) - 2R(W,Z;X,Y) - 2R(Y,Z;W,X) = 0.$$

Adding these five equalities, we get the following equation after obvious simplifications.

$$2\,R(W,X;Y,Z) - 2\,R(Y,Z;W,X) = 0,$$

and this is the identity we wanted to prove. ∎

We know from linear algebra that a symmetric bilinear form is uniquely determined by its quadratic form. More generally, when a tensor has some symmetries, it can be reconstructed from its restriction to a suitable linear subspace of its domain. For tensors having the symmetries of a curvature tensor we have the following

Proposition. Let S_1 and S_2 be tensors (or tensor fields) of valency (0,4), satisfying the following relations

$$S_i(X,Y;Z,W) = -\,S_i(Y,X;Z,W) = -\,S_i(X,Y;W,Z);$$
$$\underset{XYZ}{\circlearrowleft} S_i(X,Y;Z,W) = 0.$$

Then if $S_1(X,Y;Y,X) = S_2(X,Y;Y,X)$ for every X and Y, then $S_1=S_2$.

Proof. Consider the difference $S = S_1-S_2$. S has the same symmetries as S_1 and S_2, $S(X,Y;Y,X) = 0$ for all X,Y and we have to show $S = 0$.

We have for any X,Y,Z

$$\begin{aligned}0 = S(X,Y+Z;Y+Z,X) &= S(X,Y;Y,X) + S(X,Y;Z,X) + S(X,Z;Y,X) + S(X,Z;Z,X) = \\ &= S(X,Y;Z,X) + S(X,Z;Y,X) + \\ &\quad + (S(X,Y;Z,X) + S(Y,Z;X,X) + S(Z,X;Y,X)) \\ &= 2S(X,Y;Z,X).\end{aligned}$$

Now taking four arbitrary vectors (vector fields) X,Y,Z,W and using $S(X,Y;Z,X) \equiv 0$, we obtain

$$0 = S(X+W,Y;Z,X+W) = S(X,Y;Z,X) + S(X,Y;Z,W) + S(W,Y;Z,X) + S(W,Y;Z,W) =$$

$$= S(X,Y;Z,W) + S(W,Y;Z,X),$$

i.e., S is skew symmetric in the first and fourth variables. Thus,

$$S(X,Y;Z,W) = S(Y,X;W,Z) = - S(Z,X;W,Y) = S(Z,X;Y,W),$$

in other words, S is invariant under cyclic permutations of the first three variables. But the sum of the three equal quantities $S(X,Y;Z,W)$, $S(Y,Z;X,W)$ and $S(Z,X;Y,W)$ is 0 because of the Bianchi symmetry, thus $S(X,Y;Z,W)$ is 0. ∎

Exercise. Let S be a tensor of valency (0,4) having all the curvature tensor symmetries, and let $Q_S(X,Y) := S(X,Y;Y,X)$. Prove that $Q_S(X,Y)=Q_S(Y,X)$ and

$$\begin{aligned}6S(X,Y;Z,W) = {} & Q_S(X+W,Y+Z)-Q_S(Y+W,X+Z) + \\ & +Q_S(Y+W,X)-Q_S(X+W,Y)+Q_S(Y+W,Z)-Q_S(X+W,Z)+ \\ & +Q_S(X+Z,Y)-Q_S(Y+Z,X)+Q_S(X+Z,W)-Q_S(Y+Z,W)+ \\ & +Q_S(X,Z)-Q_S(Y,Z)+Q_S(Y,W)-Q_S(X,W).\end{aligned}$$

Definition. Let M be a Riemannian manifold, p a point on M, X and Y two non-parallel tangent vectors at p. The number

$$K(X,Y) = \frac{R(X,Y;Y,X)}{|X|^2|Y|^2-\langle X,Y\rangle^2}$$

is called the sectional curvature of M at p, in the direction of the plane spanned by the vectors X and Y in T_pM.

The name assumes that $K(X,Y)$ depends only on the plane spanned by the vectors X and Y. This is indeed so, since if $\mathbf{x}_1,\mathbf{y}_1$ and $\mathbf{x}_2,\mathbf{y}_2$ are two bases of a 2-dimensional linear space, then we can transform one of them into the other by a finite number of elementary basis transformations of the form

$$\mathbf{x} \to \alpha\mathbf{x}\ ,\ \mathbf{y} \to \beta\mathbf{y},\ \text{where } \alpha\beta\neq0;\ \mathbf{x} \to \mathbf{x}+\mathbf{y}\ ,\ \mathbf{y} \to \mathbf{y}\ ;\ \mathbf{x} \to \mathbf{y}\ ,\ \mathbf{y} \to \mathbf{x}$$

and we have the following proposition.

Proposition. If M is a Riemannian manifold with sectional curvature K, X and Y are tangent vectors at $p \in M$, α and β are non-zero scalars, then

(i) $K(X,Y) = K(X+Y,Y)$;

(ii) $K(X,Y) = K(\alpha X,\beta Y)$;

(iii) $K(X,Y) = K(Y,X)$.

Proof. (i) follows from

$$R(X+Y,Y;Y,X+Y)=R(X,Y;Y,X)+R(X,Y;Y,Y)+R(Y,Y;Y,X)+R(Y,Y;Y,Y)=R(X,Y;Y,X)$$

and

$$\begin{aligned}|X+Y|^2|Y|^2-\langle X+Y,Y\rangle^2 &= (|X|^2+|Y|^2+2\langle X,Y\rangle)|Y|^2- (\langle X,Y\rangle^2+2\langle X,Y\rangle|Y|^2+|Y|^4) = \\ &= |X|^2|Y|^2-\langle X,Y\rangle^2.\end{aligned}$$

(ii) follows from

$R(\alpha X, \beta Y; \beta Y, \alpha X) = \alpha^2 \beta^2 \, R(X,Y;Y,X)$

and

$|\alpha X|^2 |\beta Y|^2 - \langle \alpha X, \beta Y \rangle^2 = \alpha^2 \beta^2 (|X|^2 |Y|^2 - \langle X,Y \rangle^2)$.

Finally, (iii) comes from the equalities $R(X,Y;Y,X) = R(Y,X;X,Y)$ and $|X|^2|Y|^2 - \langle X,Y\rangle^2 = |Y|^2|X|^2 - \langle Y,X\rangle^2$. ▌

Definition. Riemannian manifolds, the sectional curvature function of which is constant, called *spaces of constant curvature* or simply *space forms*. The space form is *elliptic* or *spherical* if $K > 0$, K is *parabolic* or *Euclidean* if $K=0$ and is *hyperbolic* if $K < 0$.

Typical examples are the n-dimensional sphere, Euclidean space and hyperbolic space. Further examples can be obtained by factorization with fixed point free actions of discrete groups.

The following remarkable theorem sounds similarly to the theorem saying that a connected surface consisting of umbilics is contained in a sphere or plane (page 51).

Theorem (Schur). If M is a connected Riemannian manifold, $\dim M \geqslant 3$ and the sectional curvature $K(X_p, Y_p)$, $X_p, Y_p \in T_pM$ depends only on p (and does not depend on the plane spanned by X_p and Y_p, then K is constant, that is, as a matter of fact, it does not depend on p either.

Proof. By the assumption,

$$R(X,Y;Y,X) = f\,(|X|^2|Y|^2 - \langle X,Y\rangle^2)$$

for some function f. Our goal is to show that f is constant.

Consider the tensor field of valency $(0,4)$ defined by

$$S(X,Y;Z,W) = f(\langle X,W\rangle\langle Y,Z\rangle - \langle X,Z\rangle\langle Y,W\rangle).$$

It is clear from the definition that S is skew-symmetric in the first and last two arguments. S has also the Bianchi symmetry. Indeed,

$$\underset{XYZ}{\circlearrowleft} S(X,Y;Z,W) = \underset{XYZ}{\circlearrowleft} f(\langle X,W\rangle\langle Y,Z\rangle - \langle X,Z\rangle\langle Y,W\rangle) = \underset{XYZ}{\circlearrowleft} f(\langle Y,W\rangle\langle Z,X\rangle - \langle X,Z\rangle\langle Y,W\rangle) = 0.$$

We also have $R(X,Y;Y,X) = S(X,Y;Y,X)$, therefore $R = S$.

Set $\tilde{S}(X,Y;Z) = f(\langle Y,Z\rangle X - \langle X,Z\rangle Y)$. Then

$$\langle R(X,Y;Z),W\rangle = R(X,Y;Z,W) = S(X,Y;Z,W) = \langle \tilde{S}(X,Y;Z),W\rangle,$$

that is,

$$R(X,Y;Z) = \tilde{S}(X,Y;Z) \qquad \text{for all } X,Y,Z.$$

Differentiating with respect to a vector field U we get

$$(\nabla_U R)(X,Y;Z) = (\nabla_U \tilde{S})(X,Y;Z) = \nabla_U(\tilde{S}(X,Y;Z)) - \tilde{S}(\nabla_U X, Y;Z) - \tilde{S}(X, \nabla_U Y;Z) - \tilde{S}(X,Y;\nabla_U Z).$$

Since

$$\nabla_U(\tilde{S}(X,Y;Z)) = U(f)(\langle Y,Z\rangle\, X - \langle X,Z\rangle\, Y) + f\nabla_U(\langle Y,Z\rangle\, X - \langle X,Z\rangle\, Y) =$$
$$= U(f)(\langle Y,Z\rangle\, X - \langle X,Z\rangle\, Y) + f(U\langle Y,Z\rangle\, X + \langle Y,Z\rangle\nabla_U X - U\langle X,Z\rangle\, Y - \langle X,Z\rangle\nabla_U Y) =$$
$$= U(f)(\langle Y,Z\rangle X - \langle X,Z\rangle Y) +$$
$$+ f(\langle\nabla_U Y,Z\rangle X + \langle Y,\nabla_U Z\rangle X + \langle Y,Z\rangle\nabla_U X - \langle\nabla_U X,Z\rangle Y - \langle X,\nabla_U Z\rangle Y - \langle X,Z\rangle\nabla_U Y) =$$
$$= U(f)(\langle Y,Z\rangle X - \langle X,Z\rangle Y) +$$
$$+ \tilde{S}(\nabla_U X, Y; Z) + \tilde{S}(X, \nabla_U Y; Z) + \tilde{S}(X, Y; \nabla_U Z),$$

we obtain

$$(\nabla_U R)(X,Y;Z) = (\nabla_U \tilde{S})(X,Y;Z) = U(f)(\langle Y,Z\rangle X - \langle X,Z\rangle Y).$$

Using the second Bianchi identity, this gives us

$$\underset{UXY}{\circlearrowleft} U(f)(\langle Y,Z\rangle X - \langle X,Z\rangle Y) = \underset{UXY}{\circlearrowleft} (\nabla_U R)(X,Y;Z) = 0.$$

If $X \in T_pM$ is an arbitrary tangent vector to the manifold, then we can find non-zero vectors Y, $Z=U \in T_pM$ such that X,Y and U are orthogonal ($\dim M \geqslant 3$!). Then

$$0 = \underset{UXY}{\circlearrowleft} U(f)(\langle Y,Z\rangle X - \langle X,Z\rangle Y) = X(f)\langle U,U\rangle Y - Y(f)\langle U,U\rangle X.$$

Since X and Y are linearly independent, $X(f)\langle U,U\rangle = Y(f)\langle U,U\rangle = 0$. $\langle U,U\rangle$ is positive, therefore $X(f) = Y(f) = 0$, yielding that the derivative of f in an arbitrary direction X is 0. This means that f is locally constant, and since M is connected, f is constant. ∎

The curvature tensor is a complicated object containing a lot of information about the geometry of the manifold. There are some useful ways to derive some simpler tensor fields from the curvature tensor. Of course, the simplification is paid by losing information.

Definition. Let (M,∇) be a manifold with an affine connection, R be the curvature tensor of ∇. The Ricci tensor Ric of the connection is a tensor field of valency $(0,2)$ assigning to the vector fields X and Y the function $Ric(X,Y)$ the value of which at $p \in M$ is the trace of the linear mapping

$$T_pM \longrightarrow T_pM,$$
$$Z_p \longmapsto R(Z_p, X(p); Y(p)), \qquad \text{where } Z_p \in T_pM.$$

Proposition. The Ricci tensor of a Riemannian manifold is a symmetric tensor

$$Ric(X,Y) = Ric(Y,X).$$

Proof. Let $e_1,\ldots,e_n$ be an orthonormal basis in T_pM, where p is an arbitrary point in the Riemannian manifold M. We can compute the trace of a linear mapping $A:T_pM \longrightarrow T_pM$ by the formula

$$\text{trace } A = \sum_{i=1}^{n} \langle A(e_i), e_i \rangle.$$

In particular,

$$Ric(X,Y)(p) = \sum_{i=1}^{n} \langle R(e_i, X(p); Y(p)), e_i \rangle = \sum_{i=1}^{n} R(e_i, X(p); Y(p), e_i \rangle =$$

$$= \sum_{i=1}^{n} R(Y(p), e_i; e_i, X(p) \rangle = \sum_{i=1}^{n} R(e_i, Y(p); X(p), e_i \rangle = Ric(Y,X)(p). \blacksquare$$

Since the Ricci tensor of a Riemannian manifold is symmetric, it is uniquely determined by its quadratic form $X \mapsto Ric(X,X)$.

Definition. Let $X_p \in T_pM$ be a non-zero tangent vector of a Riemannian manifold M. The Ricci curvature of M at p in the direction X_p is the number

$$r(X_p) = \frac{Ric(X_p, X_p)}{|X_p|^2}.$$

Fixing an orthonormal basis $\frac{X_p}{|X_p|^2} = e_1, e_2, \ldots, e_n$ we can express the Ricci curvature as follows

$$r(X_p) = \frac{Ric(X_p, X_p)}{|X_p|^2} = \sum_{i=1}^{n} \frac{R(e_i, X_p; X_p, e_i \rangle}{|X_p|^2} = \sum_{i=2}^{n} K(X_p, e_i).$$

The meaning of this formula is that the Ricci curvature in the direction X_p is the sum of the sectional curvatures in the directions of the planes spanned by the vectors X_p and e_i, where e_i runs over an orthonormal basis of the orthogonal complement of X_p in T_pM. It is a nice geometrical corollary that this some is independent of the choice of the orthogonal basis.

With the help of a scalar product, one can associate to every bilinear function a linear transformation. For the case of the Riemannian metric and the Ricci tensor, we can find a unique $\mathcal{F}(M)$-linear transformation $\overline{Ric}: \mathfrak{X}(M) \longrightarrow \mathfrak{X}(M)$ such that

$$Ric(X,Y) = \langle X, \overline{Ric}(Y) \rangle \text{ for every } X, Y \in \mathfrak{X}(M).$$

Definition. The scalar curvature $s(p)$ of a Riemannian manifold M at a point p is the trace of the linear mapping $\overline{Ric}: T_pM \longrightarrow T_pM$.

Let us find an expression for the scalar curvature in terms of the Ricci curvature and the sectional curvature. Let $e_1, \ldots, e_n$ be an orthonormal basis in T_pM. Then

$$s(p) = \operatorname{trace} \overline{\mathrm{Ric}} = \sum_{i=1}^{n} \langle \overline{\mathrm{Ric}}(e_i), e_i \rangle = \sum_{i=1}^{n} \mathrm{Ric}(e_i, e_i) = \sum_{i=1}^{n} r(e_i),$$

i.e. $s(p)$ is the sum of Ricci curvatures in the directions of an orthogonal basis. Furthermore,

$$s(p) = \sum_{i=1}^{n} r(e_i) = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} K(e_i, e_j) = 2 \sum_{1\leqslant i<j\leqslant n}^{n} K(e_i, e_j),$$

that is, the scalar curvature is twice the sum of sectional curvatures taken in the directions of all coordinate planes of an orthonormal coordinate system in T_pM.

To finish this unit with, let us study the curvature tensor of a hypersurface M in $\mathbb{R}^n$. As we observed at the end of the previous unit, the Levi-Civita connection $\tilde{\nabla}$ of a hypersurface can be expressed as $\tilde{\nabla} = \mathbf{P}\circ\partial$, where ∂ is the derivation rule of vector fields along the hypersurface as defined in Unit 5 (page 43), $\mathbf{P}$ is the orthogonal projection of a tangent vector of $\mathbb{R}^n$ at a hypersurface point onto the tangent space of the hypersurface at that point. ∂ is essentially the Levi Civita connection of $\mathbb{R}^n$, thus, as the curvature of $\mathbb{R}^n$ is 0,

$$\partial_X\circ\partial_Y - \partial_Y\circ\partial_X = \partial_{[X,Y]}$$

for any tangential vector fields $X, Y \in \mathfrak{X}(M)$.

We have

$$\begin{aligned}\tilde{\nabla}_X\tilde{\nabla}_Y Z = \mathbf{P}(\partial_X\tilde{\nabla}_Y Z) &= \mathbf{P}(\partial_X(\partial_Y Z - \langle\partial_Y Z, \mathbf{N}\rangle\mathbf{N})) = \\ &= \mathbf{P}(\partial_X\partial_Y Z) - \mathbf{P}(X(\langle\partial_Y Z, \mathbf{N}\rangle)\,\mathbf{N}) - \mathbf{P}(\langle\partial_Y Z, \mathbf{N}\rangle\partial_X\mathbf{N}) \\ &= \mathbf{P}(\partial_X\partial_Y Z) - \langle\partial_Y Z, \mathbf{N}\rangle\partial_X\mathbf{N},\end{aligned}$$

where $X, Y, Z \in \mathfrak{X}(M)$.

Similarly,

$$\tilde{\nabla}_Y\tilde{\nabla}_X Z = \mathbf{P}(\partial_Y\partial_X Z) - \langle\partial_X Z, \mathbf{N}\rangle\partial_Y\mathbf{N}.$$

Combining these equalities with

$$\tilde{\nabla}_{[X,Y]}Z = \mathbf{P}(\partial_{[X,Y]}Z)$$

we get the following expression for the curvature tensor R of M

$$\begin{aligned}R(X,Y;Z) &= (\tilde{\nabla}_X\tilde{\nabla}_Y Z - \tilde{\nabla}_Y\tilde{\nabla}_X Z) - \tilde{\nabla}_{[X,Y]}Z = \\ &= \mathbf{P}((\partial_X\partial_Y Z - \partial_Y\partial_X Z) - \partial_{[X,Y]}Z) - \langle\partial_Y Z, \mathbf{N}\rangle\partial_X\mathbf{N} + \langle\partial_X Z, \mathbf{N}\rangle\partial_Y\mathbf{N} \\ &= \langle\partial_X Z, \mathbf{N}\rangle\partial_Y\mathbf{N} - \langle\partial_Y Z, \mathbf{N}\rangle\partial_X\mathbf{N}.\end{aligned}$$

Since $\langle Z,\mathbf{N}\rangle$ is constant zero,

$$0 = X(\langle Z,\mathbf{N}\rangle) = \langle \partial_X Z,\mathbf{N}\rangle + \langle Z,\partial_X \mathbf{N}\rangle$$

and

$$0 = Y(\langle Z,\mathbf{N}\rangle) = \langle \partial_Y Z,\mathbf{N}\rangle + \langle Z,\partial_Y \mathbf{N}\rangle.$$

Putting these equalities together we deduce that

$$R(X,Y;Z) = \langle Z,\partial_Y \mathbf{N}\rangle\partial_X \mathbf{N} - \langle Z,\partial_X \mathbf{N}\rangle\partial_Y \mathbf{N} = \langle Z,L(Y)\rangle L(X) - \langle Z,L(X)\rangle L(Y).$$

Comparing the formula

$$R(X,Y;Z) = \langle Z,L(Y)\rangle L(X) - \langle Z,L(X)\rangle L(Y)$$

relating the curvature tensor to the Weingarten map on a hypersurface with Gauss' equations proved in unit 7 we see that the curvature tensor R coincides with the curvature tensor defined there. This way, the last equation can also be considered as a coordinate free display of Gauss' equations.

Further Exercises

Exercise 12-1. Consider tensors of valency (0,4) over an n-dimensional vector space V that satisfy

$$S(X,Y;Z,W) = - S(Y,X;Z,W) = - S(X,Y;W,Z);$$
$$\underset{XYZ}{\circlearrowleft} S(X,Y;Z,W) = 0.$$

Prove that these tensors form a linear space and determine the dimension of this space.

Exercise 12-2. Prove that if X_1 and X_2 are two nonparallel principal directions at a given point p of a hypersurface M, κ_1,κ_2 are the corresponding principal curvatures, then

$$K(X_1,X_2) = \kappa_1\kappa_2.$$

What is the minimum and maximum of K(X,Y), when X and Y run over T_pM?

Exercise 12-3. Express the Ricci curvature of a hypersurface in $\mathbb{R}^{n+1}$ in a principal direction in terms of the principal curvatures.

Exercise 12-4. Express the scalar curvature of a hypersurface in terms of the principal curvatures.

Unit 13. Geodesics

Definition of geodesics, normal coordinates, variation of a curve, a variation formula for the length, description of spheres about a point with the help of normal coordinates, minimal property of geodesics.

We define the length of a smooth curve $\gamma:[a,b]\longrightarrow M$ lying on a Riemannian manifold $(M,\langle,\rangle)$ to be the integral

$$\ell(\gamma) = \int_a^b \sqrt{\langle\gamma'(t),\gamma'(t)\rangle}\, dt.$$

It is worth mentioning that the classical definition of length as the limit of the lengths of inscribed broken lines does not make sense, since the distance of points is not directly defined. The situation is just the opposite. We can define first the length of curves as a primary concept and derive from it a so called intrinsic metric $d(p,q)$, at least for connected Riemannian manifolds. $d(p,q)$ is the infimum of the lengths of all curves joining p to q. The metric enables us to define the length of "broken lines" given just by a sequence of vertices $P_1,\dots,P_N$ to be the sum of the distances between consecutive vertices. There is a theorem saying that the length of a smooth curve $\gamma:[a,b]\longrightarrow M$ is equal to the limit of the lengths of inscribed broken lines $\gamma(t_0),\gamma(t_1),\dots,\gamma(t_N)$, $a=t_0<t_1<\dots<t_N=b$ as the maximum of the distances $|t_i-t_{i-1}|$ tends to zero.

To find the analog of straight lines in the intrinsic geometry of a Riemannian manifold we have to characterize straight lines in a way that makes sense for Riemannian manifolds as well. Since the length of curves is one of the most fundamental concepts of Riemannian geometry, we can take the following characterization: a curve is a straight line if and only if for any two points on the curve, the segment of the curve bounded by the points is the shortest among curves joining the two points. A slight modification of this property could be used to distinguish a class of curves, but it is not clear whether such curves exist at all on a general Riemannian manifold.

For a physicist a straight line is the trajectory of a particle with zero acceleration or that of a light beam. This observation can also lye in the base of a definition. We only have to find a proper generalization of "acceleration" for curves lying in a Riemannian manifold. It seems quite natural to proceed as follows. The speed vectors of a curve yield a vector field along the curve. On the other hand, by the fundamental theorem of Riemannian geometry, the Riemannian metric determines a unique affine connection on the manifold which is symmetric and compatible with the metric. This connection allows us to differentiate vector fields along a curve with respect to the curve parameter. In particular, one can differentiate the speed vector field with respect to the curve parameter and may call the result the acceleration vector (or acceleration vector field along the curve).

Definition. Let M be a Riemannian manifold, γ be a curve on it and denote by $\frac{D}{dt}$ the covariant differentiation of vector fields along γ induced by the Levi-Civita connection. We say that γ is a *geodesic* if

$$\frac{D}{dt}\gamma' = 0.$$

Remark. More generally, if (M,∇) is a manifold with an affine connection, then curves satisfying $\frac{D}{dt}\gamma' = 0$ are said to be *autoparallel*. Geodesics are autoparallel curves for the Levi-Civita connection.

Proposition. The length of the speed vector of a geodesic is constant.

Proof. By the compatibility of the connection with the metric, parallel translation preserves length and angles between vectors. The definition of geodesics implies that the speed vector field is parallel along the curve, consequently consists of vectors of the same length. ∎

The proposition follows also from the equality

$$\frac{d}{dt}\langle\gamma', \gamma'\rangle = \langle\frac{D}{dt}\gamma', \gamma'\rangle + \langle\gamma', \frac{D}{dt}\gamma'\rangle = 0 .$$

As a consequence, we get that the property of "being geodesic" is not invariant under reparametrization. The parameter t of a regular geodesic is always related to the natural parameter s through an affine linear transformation i.e. $t = as + b$ for some $a,b \in \mathbb{R}$. This motivates the following definition.

Definition. A regular curve on a Riemannian manifold is a *pre-geodesic* if its natural reparameterization is geodesic.

In terms of a local coordinate system with coordinates $x_1,\dots,x_n$ a curve γ determines (and is determined by) n smooth functions $\gamma_i = x_i\circ\gamma$ $1 \leqslant i \leqslant n$. The equation $\frac{D}{dt}\gamma' = 0$ then takes the form

$$\frac{d^2\gamma_k}{dt^2} + \sum_{i,j=1}^{n} \Gamma^k_{ij}\circ\gamma \; \frac{d\gamma_i}{dt}\,\frac{d\gamma_j}{dt} = 0 \qquad \text{for } 1 \leqslant k \leqslant n .$$

The existence of geodesics depends, therefore, on the solutions of a certain system of second order differential equations.

Introducing the new functions $v_i = \frac{d\gamma_i}{dt}$ this system of n second order differential equations becomes a system of 2n first order equations

$$\begin{cases} \dfrac{d\gamma_k}{dt} = v_k \\ \dfrac{dv_k}{dt} = -\displaystyle\sum_{i,j=1}^{n} \Gamma^k_{ij}\circ\gamma \; v_i v_j \end{cases} \qquad \text{for } 1 \leqslant k \leqslant n.$$

Applying the theorem the existence and uniqueness theorem for ordinary differential equations one obtains the following.

Proposition. For any point p on a Riemannian manifold M and for any tangent vector $X \in T_pM$, there exists a unique maximal geodesic γ defined on an interval containing 0 such that $\gamma(0) = p$ and $\gamma'(0) = X$.

If the maximal geodesic through a point p with initial velocity X is defined on an interval containing $[-\varepsilon,\varepsilon]$ then there is a neighborhood U of X in the tangent bundle such that every maximal geodesic started from a point q with initial velocity $Y \in T_qM$ is defined on $[-\varepsilon,\varepsilon]$.

Since a geodesic with zero initial speed can be defined on the whole real straight line, for each point p on the manifold one can find a positive δ such that for every tangent vector $X \in T_pM$ with $\| X \| < \delta$, the geodesic defined by the conditions $\gamma(0) = p$, $\gamma'(0) = X$ can be extended to the interval $[0,1]$.

This following notation will be convenient. Let $X \in T_pM$ be a tangent vector and suppose that there exists a geodesic $\gamma : [0,1] \longrightarrow M$ satisfying the conditions $\gamma(0) = p$, $\gamma'(0) = X$. Then the point $\gamma(1) \in M$ will be denoted by $\exp_p(X)$ and called the exponential of the tangent vector X.

Using the fact that for any positive c the curve $t \longmapsto \gamma(ct)$ is also a geodesic we see that the geodesic γ is described by the formula

$$\gamma(t) = \exp_p(tX).$$

As we have observed, $\exp_p(X)$ is defined provided that $\| X \|$ is small enough. In general however, $\exp_p(X)$ is not defined for large vectors X. This motivates the following.

Definition. A Riemannian manifold is geodesically complete if $\exp_p(X)$ is defined for all $p \in M$ and all vectors $X \in T_pM$. This is clearly equivalent to the following requirement that every geodesic segment should be possible to extend to an infinite geodesic.

Proposition. For a fixed point $p \in M$, the exponential map $\exp_p$ is a smooth map from an open neighborhood of $0 \in T_pM$ into the manifold. Furthermore, the restriction of it onto a (possibly even smaller) open neighborhood of $0 \in T_pM$ is a diffeomorphism.

Proof. Differentiability of the exponential mapping follows from the theorem on the differentiable dependence on the initial point for solutions of a system of ordinary differential equations. To show that $\exp_p$ is a local diffeomorphism, we only have to show that its derivative at the point $0 \in T_pM$ is a non-singular linear mapping (see Inverse Function Theorem). Since T_pM is a linear space, its tangent space $T_0(T_pM)$ at 0 can be identified with the vector space T_pM itself. Through this identification, the derivative of the exponential map at 0 maps $T_pM \approx T_0(T_pM)$ into T_pM. We show that this derivative is just the identity map of T_pM, hence non-singular. Let X be an element of the tangent space $T_pM \approx T_0(T_pM)$. To determine where X is taken by the derivative of the exponential mapping, we represent X as the speed vector of the curve $t \longmapsto \varphi(t) = tX$ at $t = 0$. The exponential mapping takes this curve to the geodesic curve $\gamma = \exp_p \circ \varphi$ $\gamma(t) = \exp_p(tX)$, the speed vector of which at $t = 0$ is X, so the derivative of the exponential map sends X to itself and this is what we claimed. ▮

By the proposition, we can introduce a local coordinate system, based on geodesics, about each point of the manifold as follows. We fix an *orthonormal* basis in the tangent space T_pM, which gives us an isomorphism $\iota: T_pM \longrightarrow \mathbb{R}^n$ that assigns to each tangent vector its components with respect to the basis, and then take $\iota \circ \exp_p^{-1}$. The map $\iota \circ \exp_p^{-1}$ is a diffeomorphism between an open neighborhood of p and that of the origin in $\mathbb{R}^n$, therefore, it is a smooth chart on M. Coordinate systems obtained this way are called normal coordinate systems, while the inverse of them we shall call normal parametrizations .

For a Riemannian manifold M, we can define the sphere of radius r centered at $p \in M$ as the set of points $q \in M$ such that $d(p,q) = r$, where $d(p,q)$ denotes the intrinsic distance of p and q. When the radius of the sphere is increasing, the topological type of the sphere changes at certain critical values of the radius. For small radii however, the intrinsic spheres are diffeomorphic to the ordinary spheres in $\mathbb{R}^n$, and what is more, we have the following.

Theorem. The normal parameterization of a manifold about a point p maps the sphere about the origin with radius r, provided that it is contained in the domain of the parameterization, diffeomorphically onto the intrinsic sphere centered at p with radius r .

We prove the theorem later.

Definition. A variation of a smooth curve $\gamma : [a,b] \longrightarrow M$ is a smooth mapping γ_* from the rectangular domain $[-\delta,\delta] x [a,b]$ into M such that $\gamma_*(0,t) = \gamma(t)$ for all $t \in [a,b]$.

Given a variation of a curve we may introduce a one parameter family of curves γ_ε $\varepsilon \in [-\delta,\delta]$ by setting $\gamma_\varepsilon(t) = \gamma_*(\varepsilon,t)$. By our assumption these curves yield a deformation of the curve $\gamma_0 = \gamma$.

Theorem. Let γ_* be a variation of a geodesic γ. Let $\ell(\varepsilon)$ denote the length of the curve γ_ε. Then the following formula holds

$$\frac{d\ell}{d\varepsilon}(0) = < \frac{d\gamma_*}{d\ \varepsilon}(0,b),\ \frac{\gamma\ '(b)}{\|\gamma\ '(b)\|} > - < \frac{d\gamma_*}{d\ \varepsilon}(0,a),\ \frac{\gamma\ '(a)}{\|\gamma\ '(a)\|} >.$$

Proof. By the definition of the length of a curve, one has

$$\frac{d\ell}{d\varepsilon}(0) = \frac{d}{d\ \varepsilon}\int_a^b \left\| \frac{d\gamma_*}{d\ t}(\varepsilon,\tau)\right\| d\tau = \frac{d}{d\ \varepsilon}\int_a^b \sqrt{< \frac{d\gamma_*}{d\ t}(\varepsilon,\tau), \frac{d\gamma_*}{d\ t}(\varepsilon,\tau) >}\, d\tau$$

$$= \int_a^b \frac{d}{d\ \varepsilon}\sqrt{< \frac{d\gamma_*}{d\ t}(\varepsilon,\tau), \frac{d\gamma_*}{d\ t}(\varepsilon,\tau) >}\ \Bigg|_{\varepsilon=0} d\tau$$

$$= \int_a^b \frac{\frac{d}{d\ \varepsilon} < \frac{d\gamma_*}{d\ t}(\varepsilon,\tau), \frac{d\gamma_*}{d\ t}(\varepsilon,\tau) > \Big|_{\varepsilon=0}}{2\sqrt{< \frac{d\gamma_*}{d\ t}(0,\tau), \frac{d\gamma_*}{d\ t}(0,\tau) >}}\, d\tau\ .$$

With the help of the covariant differentiation induced by the Levi-Civita connection this expression can be written as follows.

$$\int_a^b \frac{\left\langle \frac{D}{d\varepsilon}\frac{d\gamma_*}{dt}(0,\tau), \frac{d\gamma_*}{dt}(0,\tau)\right\rangle}{\| \gamma'(\tau) \|} d\tau = \int_a^b \left\langle \frac{D}{d\varepsilon}\frac{d\gamma_*}{dt}(0,\tau) , \frac{\gamma'(\tau)}{\|\gamma'(\tau)\|}\right\rangle d\tau$$

By the symmetry of the connection, this is equal to

$$\int_a^b \left\langle \frac{D}{dt}\frac{d\gamma_*}{d\varepsilon}(0,\tau) , \frac{\gamma'(\tau)}{\|\gamma'(\tau)\|}\right\rangle d\tau.$$

Let us observe, that the function $t\longmapsto \left\langle \frac{d\gamma_*}{d\varepsilon}(0,t), \frac{\gamma'(t)}{\|\gamma'(t)\|}\right\rangle$ is a primitive function (=indefinite integral) for the function to be integrated. Indeed, the derivative of this function is

$$\frac{d}{dt}\left\langle \frac{d\gamma_*}{d\varepsilon}(0,t), \frac{\gamma'(t)}{\|\gamma'(t)\|}\right\rangle =$$

$$= \left\langle\frac{D}{dt}\frac{d\gamma_*}{d\varepsilon}(0,t), \frac{\gamma'(t)}{\|\gamma'(t)\|}\right\rangle + \left\langle \frac{d\gamma_*}{d\varepsilon}(0,t), \frac{D}{dt}\frac{\gamma'(t)}{\|\gamma'(t)\|}\right\rangle,$$

but the second term on the right hand side is zero since γ is geodesic. Consequently,

$$\frac{d\ell}{d\varepsilon}(0) = \int_a^b \left\langle \frac{D}{dt}\frac{d\gamma_*}{d\varepsilon}(0,\tau) , \frac{\gamma'(\tau)}{\|\gamma'(\tau)\|}\right\rangle d\tau =$$

$$= \left\langle \frac{d\gamma_*}{d\varepsilon}(0,b), \frac{\gamma'(b)}{\|\gamma'(b)\|}\right\rangle - \left\langle \frac{d\gamma_*}{d\varepsilon}(0,a), \frac{\gamma'(a)}{\|\gamma'(a)\|}\right\rangle. \blacksquare$$

Theorem. Let M be a Riemannian manifold, $p \in M$, and denote by S_r the sphere of radius r in T_pM centered at the zero tangent vector. Presume r is chosen to be so small that the exponential mapping is a diffeomorphism on a ball containing S_r and denote the exponential image of S_r by $\tilde{S}_r$. Then for any $X \in S_r$ the radial geodesic $t\longmapsto \exp_p(tX)$ is perpendicular to $\tilde{S}_r$.

Proof. Every tangent vector of $\tilde{S}_p$ can be obtained as the speed vector of a curve $\exp_p\circ\beta$ where β is a curve in S_r passing through $\beta(0) = X$. Given such a curve, let us define a variation of the geodesic $\gamma : t\longmapsto \exp_p(tX)$ in the following way

$$\gamma_*(\varepsilon,t) := \exp_p(t\beta(\varepsilon)).$$

For a fixed ε, the curve γ_ε is a geodesic of length r so $\ell(\varepsilon)$ is constant. Thus, the previous theorem implies that

$$0 = \frac{d\ell}{d\varepsilon}(0) = < \frac{d\gamma_*}{d\ \varepsilon}\ (0,1),\ \frac{\gamma\ '(1)}{\|\gamma\ '(1)\|} > - < \frac{d\gamma_*}{d\ \varepsilon}\ (0,0),\ \frac{\gamma\ '(0)}{\|\gamma\ '(0)\|} >.$$

Since $\gamma_*(\varepsilon,0) := \exp_p(0\ \beta(\varepsilon)) = p$ and , $\gamma_*(\varepsilon,1) := \exp_p(\beta(\varepsilon))$ we have $\frac{d\gamma_*}{d\ \varepsilon}(0,0) = \mathbf{0}$ and $\frac{d\gamma_*}{d\ \varepsilon}(0,1) = (\exp_p\circ\beta)'(0)$, therefore, we get

$$0 = <(\exp_p\circ\beta)'(0)\ ,\ \frac{\gamma\ '(1)}{\|\gamma\ '(1)\|} > ,$$

showing that γ intersects $\tilde{S}_r$ orthogonally. $\blacksquare$

Now we are ready to prove the theorem saying that $\tilde{S}_r$ is a sphere in the intrinsic geometry of the manifold. It is clear that for any point q on $\tilde{S}_r$ $d(p,q) \leqslant r$, since the radial geodesic from p to r has length r, so all we need is the following.

<u>Theorem</u>. If $\tilde{\gamma} : [a,b] \longrightarrow M$ is an arbitrary curve connecting p to a point of $\tilde{S}_r$, then its length is $\geqslant r$.

<u>Proof</u>. We may suppose without loss of generality that $\tilde{\gamma}(b)$ is the only intersection point of the curve with $\tilde{S}_r$ and $\tilde{\gamma}\ (t) \neq p$ for $t > a$. Then there is a unique curve γ in the tangent space T_pM such that $\tilde{\gamma} = \exp_p\circ\gamma$. Let N denote the vector field on $T_pM - \{\mathbf{0}\}$ consisting of unit vectors perpendicular to the spheres centered at the origin. N is the gradient vector field of the function $f:X \longmapsto \|\ X\ \|$ on T_pM. The theorem above shows that the exponential map takes N into a unit vector field $\tilde{N}$ on M, perpendicular to the sets $\tilde{S}_t$.

We can estimate the length of a curve as follows

$$\ell(\gamma) = \int_a^b \|\tilde{\gamma}\ '(\tau)\|\ d\tau \geqslant \int_a^b < \tilde{\gamma}\ '(\tau),\ \tilde{N}(\tilde{\gamma}(\tau)) > d\tau.$$

Since $< \tilde{\gamma}\ '(\tau),\ \tilde{N}(\tilde{\gamma}(\tau) >$ is the component parallel to $\tilde{N}(q)$ of the speed vector $\tilde{\gamma}\ '(\tau)$ with respect to the splitting $T_qM = \mathbb{R}\ \tilde{N}(q) \oplus T_q\tilde{S}_*$ at $q = \tilde{\gamma}\ (\tau)$, it is equal to the component parallel to N(X) of the speed vector $\gamma\ '(\tau)$ with respect to the splitting

$$T_X(T_pM) = \mathbb{R}\ N(X) \oplus T_XS_*$$

at $X = \gamma\ (\tau)$. Therefore,

$$< \tilde{\gamma}\ '(\tau),\ \tilde{N}(\tilde{\gamma}(\tau)) > = < \gamma\ '(\tau),\ N(\gamma(\tau)) > = < \gamma\ '(\tau),\ \text{grad}\ f(\gamma(\tau)) >$$

$$= \frac{d}{d\ t}\ f\circ\gamma\ (\tau),$$

and

$$\int_a^b < \tilde{\gamma}'(\tau), \tilde{N}(\tilde{\gamma}(\tau)) > d\tau = \int_a^b \frac{d}{dt} f\circ\gamma\ (\tau)\ d\tau = \|\gamma(b)\| - \|\gamma(a)\| = r. \blacksquare$$

The proof also shows that the equality $\ell(\gamma) = r$ holds only for curves perpendicular to the spheres $\tilde{S}_*$.

Exercise. Show that such curves are pre-geodesics.

Theorem. A smooth curve $\gamma : [a,b] \longrightarrow M$ parameterized by the natural parameter in a Riemannian manifold is geodesic if and only if there is a positive ε such that for any two values $t_1, t_2 \in [a,b]$ such that $|t_1 - t_2| < \varepsilon$ the restriction of γ onto $[t_1, t_2]$ is a curve of minimal length among curves joining $\gamma(t_1)$ to $\gamma(t_2)$.

Remark. It is not true in general, that a geodesic curve is the a curve of minimal length among curves joining the same endpoints. To see this, it is enough to consider a long arc on a great circle on the sphere.

Further Exercises

Exercise 13-1. Show that a regular curve in a hypersurface $M \subset \mathbb{R}^{n+1}$ is a geodesic if and only if its ordinary accaleration $\gamma''(t)$ is perpendicular to $T_{\gamma(t)}M$ for every t. γ is a pre-geodesic if and only if $\gamma''(t)$ is contained in the plane spanned by $\gamma'(t)$ and the normal vector of M at $\gamma(t)$.

Exercise 13-2. Show that great circles on the sphere and helices on a cylinder are pre-geodesics.

Exercise 13-3. Find a regular pre-geodesic on the cone $x^2+y^2=z^2$, different from straight lines.

Exercise 13-4. Show that straight lines on a hypersurface are pre-geodesics.

Exercise 13-5. Show that symmetry planes of a surface in $\mathbb{R}^3$ intersect the surface in pre-geodesic lines.

Exercise 13-6. Using the results of Exercise 7-2 write the differential equation of geodesics on a surface of revolution with respect to the usual parameterization. Derive from the equations *Claireaut's theorem*: For a pre-geodesic curve on a surface of revolution the quantity $d \cos\alpha$ is constant, where d denotes the distance of the curve point from the axis of symmetry, α is the angle between the speed vector of the curve and the circle of rotation passing through the curve point.