

Calculus on Manifolds

Sergei Yakovenko

The goal of this course was to teach some basics of the language today spoken by large communities of mathematicians, physicists and engineers. Accordingly, these very informal notes are by no means a substitute for the traditional textbooks. Their goal was to explain the rationale behind various geometric, algebraic and analytic constructions and how these constructions relate to each other. Using the above analogy, I tried to give examples of some typical phrases one can say in this language.

1. Introduction. Crash course on the Multivariate Calculus

1.1. Linear algebra. Field of real numbers \mathbb{R} and its completeness. Real line \mathbb{R}^1 and real spaces \mathbb{R}^n . Linear functionals, linear maps. Compositions. Rank. Group $GL(n, \mathbb{R})$, determinant. Linear algebra as a universal paradigm. Matrix algebra vs. coefficients: A as a “box” (operator) vs. $A = (a_{ij})$, the table filled by numbers. Product = composition of operators.

Warning 1.1. Other types of matrices occur: e.g., matrices of bilinear/quadratic forms. For them, multiplication is much less “justified”, if at all.

Dual space. How to plot vectors and covectors. Second dual, tensor product.

Affine space \mathbb{R}^n : no origin selected. Affine maps.

1.2. Nonlinear objects. Polynomials, polynomial maps. Domains. Functions of real variable, maps between real spaces.

Definition 1.1. Differential = “linear” (in fact, linear part of the affine) approximation of a given nonlinear map at a given point $a \in \mathbb{R}^n$.

$$f: \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad f(x) = f(a) + A(x - a) + o(|x - a|), \quad (1)$$

Date: 2016/7 academic year.

the “o small” being the term small relative to the terms explicitly written. Here A is a linear operator, a.k.a. “matrix”,

$$A = \frac{\partial f}{\partial x}(a) = f'(a) = f_*(a) = df(a) = D_a f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix} = \cdots$$

Warning 1.2. The notation of partial derivatives is ambiguous: $f_x = \frac{\partial f(x,y)}{\partial x}$ depends on which variable is chosen as y !

Definition 1.2. Differentiable maps (functions) = maps which admit approximation as above. One can use several iterative constructions to introduce maps which are differentiable several times. One (not very good) way is to require existence of iterated partial derivatives.

Smooth maps = infinitely differentiable maps (functions). All polynomials, exponentials and trigonometric functions are smooth, but there are much more of them than meets an eye.

Smoothness threads through composition according to the chain rule, $d(f \circ g)(a) = df(b) \cdot dg(a)$, where $b = g(a)$.

1.3. Tangent vectors. To be consistent with the linearity condition in (1), one has to identify the two linear spaces between the linear map A acts. These are called tangent spaces.

Definition 1.3. A tangent vector to the domain $U \subseteq \mathbb{R}^n$ at a point $a \in U$ is the pair (a, v) , where $v \in \mathbb{R}^n$ is a vector. The union of all vectors tangent to U at a is denoted by $T_a U$; it is isomorphic to \mathbb{R}^n .

The union $TU = \bigcup_{a \in U} T_a U$ is called the *tangent bundle* to U . As long as we talk about subdomains of \mathbb{R}^n , we obviously have $TU = U \times \mathbb{R}^n$, that is, an open subspace of \mathbb{R}^{2n} , but for general manifolds it is not true.

One could think of a tangent vector as a “free” vector from \mathbb{R}^n pinned down at the point a . Note that two tangent spaces at two different points are disjoint.

Using this notion, the difference $x - a$ in (1) can be identified with a vector v from $T_a U$. In the same vein the image Av is identified with the vector tangent to $T_b V$, $b = f(a)$. This is the convention that will be always in place.

Definition 1.4. Diffeomorphism $f: U \rightarrow V$ for $U, V \subset \mathbb{R}^n$ is a one-to-one smooth map whose set theoretic inverse $f^{-1}: V \rightarrow U$ is also smooth.

Theorem 1.1 (Inverse function theorem). *If f is a smooth map and its differential at an interior point $a \in U$ is invertible, then there is a small neighborhood $(U, a) \subseteq \mathbb{R}^n$ such that f is a diffeomorphism between this neighborhood and its image.*

Proof. Assume that the differential is an identity matrix and $a = 0$, so $f(x) = x + h(x)$, where h is a vector function with all partial derivatives vanishing at the origin. We look for solution of the equation $y = x + h(x)$ in the form $x = y + g(y)$. Plugging it into the equation, we see $y = y + g(y) + h(y + g(y))$, that is, $g(\cdot) = -h(\cdot + g(\cdot))$.

Thus g is a fix point for an operator $\psi \mapsto -h \circ (\text{id} + \psi)$ on a suitable space of functions (that has to be properly described, of course). This operator is *very strongly contracting* on the space of smooth functions defined on a sufficiently small neighborhood of the origin. \square

1.4. Dimension one: a notable exception. The strategic difference between functions and maps: \mathbb{R} has the structure of a field (and also one-dimensional linear space over itself). In general, \mathbb{R}^n has no extra algebraic structure except for that of a linear \mathbb{R} -space.

Notable exceptions: $n = 2$: $\mathbb{R}^2 \simeq \mathbb{C}$ is also a field; $n = 4$: \mathbb{R}^4 is a non-commutative body of quaternions; $n = 8$: non-associative Caley's octonions.

For metaphysical purposes we will sometimes try to distinguish explicitly between the real line \mathbb{R}^1 as a linear one-dimensional space over the field \mathbb{R} and the field itself. In other words, elements of \mathbb{R}^1 are forbidden to multiply between themselves, only by elements from \mathbb{R} . To stress the difference, we will sometimes use the notation $\mathbb{R}_{\text{field}}$. In particular, $\mathbb{R}_{\text{field}}$ contains a distinguished element 1 (the multiplicative unity) which will be denoted by $\mathbf{1}$.

Warning 1.3. The linear map between two one-dimensional spaces in coordinates is a (1×1) -matrix with a single element. Such matrix can be identified with the numeric value of this element. This allows to say that the differential of a smooth map $f: \mathbb{R}^1 \rightarrow \mathbb{R}^1$ at a point a “is” the number, called the derivative of f at a and denoted $f'(a)$. This identification provides the context in which the derivative of such a function is again a function $f': a \mapsto f'(a)$ of the same type, hence the operation of derivation can be iterated, giving rise to higher derivatives. This is important as one can use this construction to define inductively functions which are differentiable any finite number of times, infinitely-differentiable functions etc. Doing this for maps between higher-dimensional domains is much more technically involved.

1.5. Smooth functions. Let $U \subseteq \mathbb{R}^n$ be an open domain (connected open set). We can consider all real-valued *smooth* (= infinitely differentiable) functions in U , which constitute an \mathbb{R} -algebra, i.e., a linear (infinite-dimensional) topological space $\mathcal{C}(U)$ over the field \mathbb{R} with the multiplication operation (sometimes we will use the more verbose notation $C^\infty(U)$ instead of $\mathcal{C}(U)$ or just \mathcal{C}).

It is the special property of functions: since $\mathbb{R} = \mathbb{R}_{\text{field}}$ is a field, functions can be not only added between themselves, but also multiplied (and sometimes divided).

Remark 1.1. For a smooth function its differential df at a point $a \in U$ is a linear map from $T_a U \simeq \mathbb{R}^n$ to $T_{f(a)} \mathbb{R}^1 \simeq \mathbb{R}^1$. If we identify (tautologically) \mathbb{R}^1 with $\mathbb{R}_{\text{field}}$, then the differential becomes an element of the *dual* vector space $T_a^* U \simeq (\mathbb{R}^n)^*$.

1.6. Curves. A curve (more precisely, a parameterized curve) is a map $\gamma: U \rightarrow V \subseteq \mathbb{R}^n$, where $U \subseteq \mathbb{R}^1$ is an open connected subset of the real line. The image $\gamma(U)$ is the non-parameterized curve. The differential is a linear map $d\gamma(a): T_a U \rightarrow T_{\gamma(a)} V$. The image $v(a) = d\gamma(a)\mathbf{1}$ is a vector $v \in T_{\gamma(a)} V \simeq \mathbb{R}^n$, called the velocity vector at the point $b = \gamma(a)$.

We will often interpret the coordinate on \mathbb{R}^1 as time and denote the velocity by the dot, $v(t) = \dot{\gamma}(t)$.

If $\dot{\gamma}(a) = v \in T_a U$, we say that the curve γ is tangent to the vector v at the point a .

2. Vector fields and what they are good for

2.1. Vector fields and ordinary differential equations. A vector field in a domain $U \subseteq \mathbb{R}^n$ is a (smooth) map $X: U \rightarrow TU$ such that $X(a) \in T_a U$ for all $a \in U$. Speaking plainly, this is a collection of vectors attached to all points of U , which vary differentiably (smoothly) on the point. The set of all vector fields will be denoted by $\mathcal{X}(U)$.

Given a vector field, one can look for:

- a curve tangent to $v = X(a)$ at a given point $a \in U$;
- a curve, tangent to all vectors $v(\gamma(t))$ at each moment t ;
- a family of curves tangent to all vectors at all points in U .

The first task is elementary: it suffices to take the curve $\gamma(t) = a + tv$, $a \in U$, $v = X(a) \in \mathbb{R}^n$. The second task amounts to finding solution to a system of ordinary differential equations with a specified initial condition.

Theorem 2.1 (Main existence and uniqueness theorem for ODEs). *If $X \in \mathcal{X}(U)$ is a smooth vector field, then for any point $a \in U$ there exists a unique smooth curve $\gamma_a: (\mathbb{R}, 0) \rightarrow (U, a)$ which is tangent to the vector $X(\gamma(t))$ at each point $\gamma(t)$, $t \in (\mathbb{R}, 0)$, and satisfies the “initial condition” $\gamma(0) = a$.*

This curve is unique in the sense that any other solution $\tilde{\gamma}_a$ coincides with γ_a on a sufficiently small neighborhood of the origin.

As a function of a and t , the curve γ is smooth.

Proof. In coordinates $x = (x_1, \dots, x_n)$ the vector field X is given by a tuple of smooth functions $(v_1(x), \dots, v_n(x))$, and a curve $\gamma: t \mapsto x(t)$ tangent to X by a tuple of functions $(x_1(t), \dots, x_n(t))$. The tangency condition is expressed as the system of identities

$$\frac{dx_i(t)}{dt} = v_i(x_1(t), \dots, x_n(t)), \quad x(0) = a, \quad i = 1, \dots, n.$$

Now go to your favorite textbook on ODE’s or apply the contracting map principle to the Picard approximations which send an arbitrary smooth curve $x(\cdot)$ such that $x(0) = a$ to the new curve

$$y(t) = a + \int_0^t v(x(\tau)) d\tau, \quad t \in I = [-\varepsilon, \varepsilon] \subseteq \mathbb{R}^1.$$

(Prove that this Picard map is contracting if $\varepsilon > 0$ is small enough). \square

Warning 2.1. If U is bounded, then any solution can be extended to a curve which crosses the boundary of U . For unbounded domains the domain of definition of γ may be smaller than \mathbb{R} (explosion of solutions). For instance, if $v(x) = x^2$, then the solution with $x(0) = 1$ exists only on finite interval. Find it by solving the equation $\frac{dx}{dt} = x^2$ explicitly.

If X is not smooth enough, solution can be non-unique. For instance, the equation $\dot{x} = \sqrt{|x|}$ on \mathbb{R}^1 has at least two solutions with $x(0) = 0$, which vanish identically for $t \leq 0$. One is the identical zero, the other is positive, $x(t) = \frac{1}{2}t^2$ for $t > 0$. Find all solutions and describe their smoothness.

If we replace a field $X \in \mathcal{X}(U)$ by the field $\tilde{X} = \lambda X$, $\lambda \in \mathbb{R}_{\text{field}}$, then one can easily integrate the new equation by setting $\tilde{\gamma}_a(t) = \gamma_a(\lambda t)$. In the more general case $\tilde{X} = \varphi X$, $\varphi \in \mathcal{C}(U)$, one can also find solutions by an integration. In particular, if $\varphi > 0$, then the images of the curves γ_a and $\tilde{\gamma}_a$ locally coincide.

2.2. Flows of vector fields. One can consider simultaneously all integral curves of a vector field $X \in \mathcal{X}(U)$ through all points of U . It is a

map of the (subset of the) Cartesian product $\mathbb{R} \times U$ into U , which sends the pair (t, a) into the point $\gamma_a(t)$ of the curve γ_a which satisfies the initial condition $\gamma_a(0) = a$. (We assume that such curve can be extended to the point $t \in \mathbb{R}$). We will denote this map $F_X: (t, a) \mapsto F_X^t(a)$ for the following reason.

Proposition 2.1. Assume that both $b = F_X^t(a)$ and $c = F_X^s(b)$ are defined for some $a \in U$ and $t, s \in \mathbb{R}$. Then $F_X^{t+s}(a)$ is also defined, and

$$F_X^s(F_X^t(a)) = F_X^{t+s}(a).$$

We will often omit the indication of X if the field is clear from the context.

Proof. It follows from the existence and uniqueness of integral trajectories. \square

Assuming that the field X and domain U are so nice that the flow F^t is defined for all $t \in \mathbb{R}$ as a smooth self-map of U , the above proposition means that

$$F^t \circ F^s = F^{s+t} = F^s \circ F^t, \quad F^0 = \text{id}.$$

In other words, the collection $\{F_X^t : t \in \mathbb{R}\}$ is a commutative subgroup of the group $\text{Aut}(U)$ of smooth self-diffeomorphisms of U .

The map $F_X: \mathbb{R} \times U \rightarrow U$ is smooth. Its differential can be immediately computed at all points $(0, a)$: it takes the vector $(\mathbf{1}, 0) \in T_{(1,a)}\mathbb{R} \times U$ to $v = X(a) \in T_aU$ and all vectors $(0, v)$ into v (“identically”).

2.3. Rectification theorem. Assume that the vector $v = X(a) \in T_aU \simeq \mathbb{R}^n$ is nonzero. Then one can find a complementary $(n - 1)$ -dimensional subspace $\Pi \subset T_aU$ such that $v \notin \Pi$.

Consider the restriction of the flow map F_X on $\mathbb{R} \times a + \Pi \subset \mathbb{R} \times U$. This is a smooth map between two n -dimensional spaces. The above computation shows that the differential of this restriction is non-degenerate (in particular, surjective) linear map.

By the Inverse Function Theorem this map is locally invertible. This means that there exists a small neighborhood (U, a) such that every point x from this neighborhood is representable as $F^t(y)$, $t \in (\mathbb{R}^1, 0)$, $y \in \Pi \simeq \mathbb{R}^{n-1}$.

If we use (t, y) as the new coordinates near U , then in these coordinates the flow map takes the simplest form $F^s(t, y) = (s + t, y)$. Trajectories are straight lines parallel to the t -axis. The corresponding vector field is given by the vector-function $(1, 0, \dots, 0)$ which is constant (independent of the point).

2.4. Canonical representation. Algebras and morphisms. It is very convenient in parallel with the geometric study of maps, curves, vector fields etc., to study infinite-dimensional but *linear* objects.

The main object associated with a domain $U \subseteq \mathbb{R}^n$ is the *commutative \mathbb{R} -algebra* $\mathcal{C}(U)$ of C^∞ -smooth functions in U . This algebra with the \mathbb{R} -linear operations, the pointwise multiplication of functions, can be equipped by several topologies (depending on whether U is bounded or not).

If $F: U \rightarrow V$ is a smooth map, then it defines a map of algebras $F^*: \mathcal{C}(V) \rightarrow \mathcal{C}(U)$ (note the direction change) which is a continuous homomorphism of algebras. The isomorphism takes $g \in \mathcal{C}(V)$ into $F^*g = g \circ F \in \mathcal{C}(U)$.

The correspondence $F \mapsto F^*$ is *functorial*, in particular, if $F: U \rightarrow V$ and $G: V \rightarrow W$, so that $G \circ F: U \rightarrow W$, then $(G \circ F)^*: \mathcal{C}(W) \rightarrow \mathcal{C}(U)$ is the composition, $(G \circ F)^* = F^*G^*$ (traditionally, for *linear* operators we omit the sign of composition \circ , replacing it by the null symbol for “matrix” multiplication). Note the inversion of the order!

Consequently, if F is a diffeomorphism between U and V , then F^* is an isomorphism of algebras. In this case, $(F^{-1})^* = (F^*)^{-1}$.

Conversely, *any* continuous homomorphism $A: \mathcal{C}(V) \rightarrow \mathcal{C}(U)$ is induced by some smooth map F , so that $A = F^*$.

Proof. Consider the coordinate functions $y_i \in \mathcal{C}(V)$ and look at the map $F = (f_1, \dots, f_n)$, where $f_i = Ay_i$. By homomorphism, $Au = F^*u$ for all polynomials $u \in \mathbb{R}[y_1, \dots, y_n]$. Since polynomials are dense, $F = f^*$ on all functions. \square

2.5. Vector fields as derivations. If $U \subseteq \mathbb{R}^n$ and $X \in \mathcal{X}(U)$ is a smooth vector field, then any function $\varphi: U \rightarrow \mathbb{R}$ can be differentiated along X producing a new function, denoted sometimes $X\varphi$, sometimes $L_X\varphi \in \mathcal{C}(U)$.

Definition 2.1. The Lie derivative of a smooth function $\varphi \in \mathcal{C}(U)$ along $X \in \mathcal{X}(U)$ is a function $X\varphi$ whose value at any point $a \in U$ is the derivative

$$X\varphi(a) = \left. \frac{d}{dt} \right|_{t=0} \varphi(F_X^t(a))$$

Using the chain rule of derivation, we compute derivatives along the curve $\gamma(t) = F_X^t(a)$:

$$\left. \frac{d}{dt} \right|_{t=0} \varphi(\gamma(t)) = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i}(a) \cdot \left. \frac{dx_i(t)}{dt} \right|_{t=0} = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i}(a) v_i(a).$$

This implies that:

- The value $X\varphi(a)$ depends only on the vector $v = X(a)$, and depends in a linear way;
- As an operator $X: \mathcal{C}(U) \rightarrow \mathcal{C}(U)$, is an \mathbb{R} -linear operator which satisfies the *Leibniz rule*,

$$X(\varphi\psi) = \varphi \cdot (X\psi) + \psi \cdot (X\varphi). \quad (2)$$

Conversely, if $L: \mathcal{C} \rightarrow \mathcal{C}$ is an \mathbb{R} -linear self-map of \mathcal{C} to itself, which satisfies the Leibniz rule $L(\varphi\psi) = \varphi \cdot L\psi + \psi \cdot L\varphi$, then L is a Lie derivative along a suitable vector field X .

Indeed, by the Leibniz rule $L1 = L(1 \cdot 1) = 2L1$, so $L1 = 0$ and by linearity $L\lambda = 0$ for any constant $\lambda \in \mathbb{R}$.

Let $v_i(x)$ be the L -images of the coordinate functions, $v_i = Lx_i$, $x_i \in \mathcal{C}$. We show that the vector field X with such coordinates indeed acts on an arbitrary function φ at an arbitrary point as L acts, $L\varphi(a) = \sum v_i(a) \frac{\partial \varphi}{\partial x_i}(a)$.

By the Hadamard formula which allows to represent each smooth function φ , say, near the origin under the form

$$\varphi(x) = \varphi(a) + \sum_{i=1}^n (x_i - a_i) \varphi_i(x), \quad \varphi_i \in \mathcal{C}, \quad \varphi_i(a) = \frac{\partial \varphi}{\partial x_i}(a).$$

Applying L to this expression and using again the Leibniz rule, we conclude that

$$(L\varphi)(a) = 0 + \sum_{i=1}^n \left((Lx_i(x) - 0) \varphi_i(a) + 0 \cdot (L\varphi_i)(a) \right) = (X\varphi)(a).$$

Lie derivations (first order differential operators on $\mathcal{C} = \mathcal{C}(U)$, satisfying the Leibniz rule) form a *module* over the algebra \mathcal{C} : they can be added between themselves and multiplied by functions from \mathcal{C} . In a given coordinate system the basis of this module can be chosen as the (standard partial) derivations $X_i = \frac{\partial}{\partial x_i}$ which act on the coordinate functions in the predictable way, $X_i x_j = \delta_{ij}$. Any other vector field can be then expanded as $X = \sum v_i X_i = \sum x_i \frac{\partial}{\partial x_i}$. This is a standard practice to denote vector fields.

2.6. Commutator. Although derivations from $\mathcal{X}(U)$ take the algebra $\mathcal{C}(U)$ into itself as \mathbb{R} -linear operators and hence can be composed with each other, the composition in general will not satisfy the Leibniz rule. Indeed, for $X, Y \in \mathcal{X}(U)$ and $\varphi, \psi \in \mathcal{C}(U)$, we have

$$XY(\varphi\psi) = X(\varphi \cdot Y\psi + \psi \cdot Y\varphi) = X\varphi \cdot Y\psi + \varphi \cdot XY\psi + X\psi \cdot Y\varphi + \psi \cdot XY\varphi.$$

The terms $X\varphi \cdot Y\psi + X\psi \cdot Y\varphi$ thwart the Leibniz rule for XY . However, if we consider the difference $XY - YX$, the obstructing terms will cancel

each other, and the operator $Z = XY - YX$ is again a derivation (i.e., an operator of the first rather than second order).

Definition 2.2. The *commutator* of two vector fields $X, Y \in \mathcal{X}(U)$ is the vector field associated with the derivation $[X, Y] = XY - YX$.

Remark 2.1. The same fact can be derived from the first year calculus assertion (“Clairaut theorem” or “Schwarz theorem”) that the second partial derivatives do not depend on the order, so that

$$\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right] = 0.$$

2.7. Flows and fields in the canonical representation. The relationship between a field $X \in \mathcal{X}(U)$ and the respective flow $\{F_X^t \in \text{Aut}(U) : t \in \mathbb{R}\}$ is expressed by the formula

$$\left. \frac{dF_X^t(a)}{dt} \right|_{t=0} = X(a) \in T_a U,$$

the left hand side being interpreted as the velocity vector of the smooth curve $t \mapsto F_X^t(a)$ at the point a , the right hand side as the tangent vector to U at a .

Assume we are guaranteed that the flow is well defined as a family of self-maps of U and hence we have a one-parametric family of automorphisms $\{(F_X^t)^* : t \in \mathbb{R}\}$ of the algebra $\mathcal{C}(U)$. Then for any test function $\varphi \in \mathcal{C}(U)$ we have a one-parametric family of functions $\{\varphi_t = (F_X^t)^* \varphi : t \in \mathbb{R}\}$ and can compute the (partial) derivative $\left. \frac{d\varphi_t}{dt} \right|_{t=0}$. The answer is not surprising: the derivative is equal to $X\varphi$. In other words, we have the identity between the linear operators on $\mathcal{C}(U)$,

$$\left. \frac{d}{dt} \right|_{t=0} (F_X^t)^* = X.$$

If the algebra $\mathcal{C}(U)$ were finite-dimensional, we would have a one-parametric group of linear operators $\{B^t\}$ such that $\left. \frac{d}{dt} \right|_{t=0} = A$ is a linear operator called the *generator*. The well-known theory of linear systems provides an explicit formula for recovering the family $\{B^t\}$ from the given operator A :

$$B^t = \exp(tA) = E + tA + \frac{t^2}{2!}A^2 + \cdots + \frac{t^n}{n!}A^n + \cdots$$

Leaving aside the issue of convergence of the series, this formula justifies the notation

$$(F^t)^* = \exp tX$$

as the identity between the operators on $\mathcal{C}(U)$.

Example 2.1. Assume that $U = \mathbb{R}^1$ and $X = \frac{\partial}{\partial x}$. The corresponding differential equation is trivial, $\dot{x} = 1$. The flow map is also easy to compute: $(F^t)^*\varphi(x) = \varphi(x + t)$. The above identity then takes the form

$$\varphi(x + t) = \varphi(x) + t\varphi'(x) + \frac{t^2}{2!}\varphi''(x) + \cdots + \frac{t^n}{n!}\varphi^{(n)}(x) + \cdots .$$

It holds for all analytic functions $\varphi \in \mathcal{C}(U)$, but in general fails. Yet the mnemonic helps anyway and may be turned into meaningful computations if we replace functions by their finite order jets at a given point.

2.8. Action of diffeomorphisms on vector fields. Let $F: U \rightarrow V$ be a smooth map. How it acts on vector fields?

Geometric push-forward. Suppose $X \in \mathcal{X}(U)$ is a vector field on U . Take an arbitrary point $a \in U$ and consider a curve $\gamma_a: (\mathbb{R}^1, 0) \rightarrow (U, a)$ passing through a and tangent to $v(a)$. The push-forward $\eta_b = f_*\gamma_a$ will be a smooth curve in V passing through $b = f(a)$. We can attach its velocity vector $w = \dot{\eta}_b$ to the point $b \in V$ and call it the “push-forward” of $v = X(a)$ by F . Bad news: if F is not surjective, then the push-forward is not everywhere defined, and if the preimage $F^{-1}(b)$ consists of more than one point, then more than one vector can be obtained this way.

Easy computation shows that $w = dF(a) \cdot v$.

Algebraic push-forward. We need to construct a Lie derivation of $\mathcal{C}(V)$ which becomes the Lie derivative X on $\mathcal{C}(U)$ after the pull-back F^* . This is straightforward: for any $\varphi \in \mathcal{C}(V)$ consider $\psi = F^*\varphi$ and apply X . The resulting operator $\mathcal{C}(V) \rightarrow \mathcal{C}(U)$ will satisfy the Leibniz rule as required, but we need to get back to $\mathcal{C}(V)$. In general it is impossible, but if F is a diffeomorphism, then we can use its inverse F^{-1} to do the job and define $Y: \mathcal{C}(V) \rightarrow \mathcal{C}(V)$, $Y = (F^{-1})^*XF^*$. As it satisfies the Leibniz rule, Y is the Lie derivative associated with a vector field on V .

The push-forward Y of the vector field X by a diffeomorphism F is denoted $Y = F_*X$. By construction, $F^*Y = XF^*$ as operators $\mathcal{C}(V) \rightarrow \mathcal{C}(U)$. The two Lie derivations are conjugated by an isomorphism F^* in the “linear algebraic” sense. Moreover, the latter equation may be satisfied even for non-invertible smooth maps, if X is special enough: in this case the two vector fields are said to be F -related.

Example 2.2. Think of a projection $F: \mathbb{R}^2 \rightarrow \mathbb{R}^1$ on the horizontal axis and a vector field X whose horizontal component is independent of the vertical coordinate.

2.9. Local nature of homomorphisms and derivations. The rings $\mathcal{C}(U)$, as contrasted to rings of polynomial functions, contain cutoff functions. A cutoff function for a point $a \in U$ by definition is a function χ that is identically equal to 1 in some neighborhood (U, a) and vanishes identically outside the twice bigger neighborhood.

Theorem 2.2. *If L is a Lie derivation of $\mathcal{C}(U)$ associated with a vector field X , then for any two functions φ, ψ which coincide in some open neighborhood of a point $a \in U$, the functions $L\varphi$ and $L\psi$ also coincide in a small neighborhood of a . In other words, values of $L\varphi$ near a do not depend on the behavior of φ away from a .*

Proof. The difference $\varphi - \psi$ vanishes in a neighborhood of a . Let χ be a cutoff function which is $\equiv 1$ near a , $\equiv 0$ outside and such that $\chi(\varphi - \psi) \equiv 0$ everywhere. Applying L we conclude that $0 \equiv (L\chi)(\varphi - \psi) + \chi(L\varphi - L\psi)$. Since the first term is zero near a and χ is one there, we get vanishing $L\varphi - L\psi$. \square

Remark 2.2. This construction is a cheap trick that allows to modify a vector field $X \in \mathcal{X}(U)$ so that it will be unchanged near each specific point $a \in U$ but the modified vector field will have the flow map to be defined as a self-map of U , thus avoiding any problems with the domain of definition.

It suffices to replace X by the field χX , where $\chi \in \mathcal{C}(U)$ is a smooth function identically equal to 1 near a and vanishes identically outside some (larger) neighborhood of a .

3. Manifolds

Morally, manifolds are topological spaces which locally look like open balls of the Euclidean space \mathbb{R}^n .

One can construct them by piecing together such balls (“cells”) using smooth maps identifying points in the overlaps.

Example 3.1. Circle, cylinder, torus.

Cube made of its unfolding. Appearance of corners.

Don’t forget the flaps!

However, it is technically more convenient to proceed in the opposite direction.

3.1. Formal definition.

Definition 3.1. A (differentiable) manifold M is a Hausdorff second-countable topological space equipped with a family (“atlas”) of mutually agreeing charts covering the whole of M .

A chart is a homeomorphism $x: U \rightarrow \mathbb{R}^n$ from an open domain $U \subseteq M$ and a ball in the Euclidean space \mathbb{R}^n . The value $x(a) = (x_1(a), \dots, x_n(a))$ is called as *coordinates* of the point $a \in M$

Two charts $x: U \rightarrow \mathbb{R}^n$ and $y: V \rightarrow \mathbb{R}^n$ defined on the overlapping domains $U \cap V \neq \emptyset$ are said to agree on the intersection, if the *transition map* $y \circ x^{-1}: x(U \cap V) \rightarrow \mathbb{R}^n$ is a C^∞ -smooth diffeomorphism (between domains in \mathbb{R}^n), ditto $x \circ y^{-1}$.

The number n (necessarily the same for all charts) is called the *dimension* of the manifold.

Warning 3.1. Formally, the same space can be covered by charts in many different ways, producing different atlases. However, the corresponding *smooth structures* may coincide. This happens when charts from different atlases can be shuffled together without creating disagreements. To play safe, one should take the maximal (in the Zorn lemma sense) atlas of mutually agreeing charts.

In practice we will always deal with finite or at most countable atlases.

Remark 3.1. The topological constraints are imposed to exclude pathological examples like a line with two zeros, huge spaces etc.

3.2. Examples. Euclidean spaces and their open domains with atlases of trivial charts.

Subspaces of the Euclidean spaces of smaller dimension (via the Implicit Function theorem). Spheres.

Quotient spaces. Tori. Projective spaces.

Linear groups. Orthogonal and Special Orthogonal groups.

Cartesian products.

Configuration spaces of mechanical systems.

What is not a manifold? Topological and analytic obstructions.

4. Constructions on manifolds

4.1. Chart-wise approach. Assume that we are given a smooth manifold with an atlas of charts $x_\alpha: U_\alpha \rightarrow \mathbb{R}^n$, $\alpha \in \mathfrak{A}$ (the atlas may be finite or even uncountable, what’s important is that it is self-consistent). Let $F_{\alpha\beta} = x_\beta \circ x_\alpha^{-1}$ the transition maps, $\alpha, \beta \in \mathfrak{A}$.

Definition 4.1. A function $f: M \rightarrow \mathbb{R}$ is called C^∞ -smooth (or simply smooth) on M , if all compositions $f_\alpha: f \circ x_\alpha^{-1}$, $\alpha \in \mathfrak{A}$, are C^∞ -smooth on the domains $D_\alpha = x_\alpha(U_\alpha) \subseteq \mathbb{R}^n$.

We write then that $f \in \mathcal{C}(M)$.

Any function $f \in \mathcal{C}(M)$ can be thus be “recorded” by a collection of functions $\{f_\alpha(x_\alpha)\}_{\alpha \in \mathfrak{A}}$, defined in the domains $x_\alpha \in D_\alpha$, and related by the appropriate identities in $D_{\alpha\beta} = x_\alpha(U_\alpha \cap U_\beta)$ and $D_{\beta\alpha} = x_\beta(U_\alpha \cap U_\beta)$:

$$f_\alpha \circ F_{\beta\alpha} = f_\beta \quad \text{on} \quad D_{\alpha\beta} \quad \forall \alpha, \beta \in \mathfrak{A}.$$

The analogous definition for tangent vectors and vector fields has to be even less intuitive. Unlike the abstract notion of a real-valued function which can be defined on any set, there is no abstract notion of a vector tangent to a topological space, thus the key element in the construction disappears. The best one can do is an artificial construction as follows.

Definition 4.2. A tangent vector to a manifold M at a point $a \in M$ is a collection of vectors $\{v_\alpha \in \mathbb{R}^n: \alpha \in \mathfrak{A}\}$ attached at the points $x_\alpha(a) \in D_\alpha$, which are agreeing between themselves in the following sense:

$$v_\beta = dF_{\alpha\beta}(x_\alpha(a)) \cdot v_\alpha, \quad \text{if} \quad a \in D_{\alpha\beta}.$$

The set of all tangent vectors at a point a is denoted by $T_a M$.

Each chart x_α provides an isomorphism between $T_a M$, $a \in M$, and \mathbb{R}^n . The (disjoint) union of all tangent spaces is denoted TM . The above formula is in fact the hidden definition of the structure of a smooth manifold on TM . It consists of the atlas of charts on $U_\alpha \times TU_\alpha$ with the transition functions given by the pairs

$$(x_\alpha, v_\alpha) \mapsto (x_\beta, v_\beta), \quad x_\beta = F_{\alpha\beta}(x_\alpha), \quad v_\beta(x_\beta) = (dF_{\alpha\beta})(x_\alpha) \cdot v_\alpha(x_\alpha).$$

Definition 4.3. A vector field on a manifold M is a smooth map $X: M \rightarrow TM$ such that $X(a) \in T_a M$.

One can associate with any $X \in \mathcal{X}(M)$ a collection $\{X_\alpha \in D_\alpha\}_{\alpha \in \mathfrak{A}}$ of vector fields in D_α , which are related by the corresponding formulas.

This line of definitions leads to the idea of a *tensor*, generalizing that of a tangent vector. The tensors are “collection of numbers associated with each chart. which are transformed according to certain rules”. Axiomatizing the computation rules is a cheap way to circumvent the inconvenient questions about the nature of the things. “The method of postulating what we want has many advantages; they are the same as the advantages of theft over honest toil” (Bertrand Russell, 1919).

4.2. Geometric attempt. Another attempt to define the notion of a tangent vector is via smooth curves on the manifold.

Definition 4.4. A map $\gamma: (\mathbb{R}, 0) \rightarrow M$ is a smooth (parameterized) curve on a manifold M , if its images $\gamma_\alpha = x_\alpha \circ \gamma$ are smooth curves in D_α for all $\alpha \in \mathfrak{A}$.

Two curves γ, η are mutually tangent at a point $t = 0$, if $\|\gamma_\alpha(t) - \eta_\alpha(t)\| = o(|t|)$ for some (hence, for all) charts covering a .

Definition 4.5. A tangent vector at a point $a \in M$ is the equivalence class of curves as above.

Looking at the same curve at different charts, one can verify that this definition leads to the same object(s). A vector field as a family of tangent vectors smoothly dependent on the attachment point can be easily formalized.

However, this approach makes it highly non-obvious why the collection of all tangent vectors at the same point has the structure of a linear (vector) space. This can be proved by computation, OK, but remember about of “theft vs. toil”.

4.3. Algebraic approach. It seems that the most aesthetically appealing approach is mimicking that of algebraic geometry. The only object that has to be defined chart-wise is that of the ring (\mathbb{R} -algebra) $\mathcal{C}(M)$ of C^∞ -functions.

Other derived objects can be defined in terms of this ring.

Problem 4.1. Consider the subalgebra of $C(\mathbb{R})$ which consists of all continuous functions which are C^∞ -smooth outside the origin $x = 0$, all derivatives have one-sided limits as $x \rightarrow \pm 0$, but these limits can be different from two sides.

Show that this ring is non-isomorphic (in the algebraic sense) to the ring $\mathcal{C}(\mathbb{R}) = C^\infty(\mathbb{R})$.

Hint: assuming such an isomorphism exists, try to guess what can be the image of $f(x) = |x|$.

Definition 4.6. A point map is a morphism $a: \mathcal{C}(M) \rightarrow \mathbb{R}$, $f \mapsto f(a)$. It can be associated with a maximal ideal $I_a \in \mathcal{C}(M)$ of functions, $I_a = \{f \in \mathcal{C}(M): f(a) = 0\}$.

A smooth map $F: M \rightarrow N$ is the point map associated with a continuous morphism of the algebras $\mathcal{C}(N) \rightarrow \mathcal{C}(M)$. In other words, F is smooth, if the pullback $F^*: \mathcal{C}(N) \rightarrow \mathcal{C}(M)$ takes $\mathcal{C}(N) \subsetneq C(N)$ to $\mathcal{C}(M) \subsetneq C(M)$.

A tangent vector is an \mathbb{R} -linear map $v_a: \mathcal{C}(M) \rightarrow \mathbb{R}$, which satisfies the Leibniz formula

$$v_a(fg) = f(a)v_a(g) + g(a)v_a(f).$$

Respectively, a vector field X is an \mathbb{R} -linear self-map of $\mathcal{C}(M)$ which satisfies the Leibniz rule

$$X(fg) = f \cdot Xg + g \cdot Xf \quad \forall f, g \in \mathcal{C}(M).$$

A smooth curve through a point a is a smooth map $\gamma: (\mathbb{R}^1, 0) \rightarrow M$. Its velocity vector at a is naturally a tangent vector $\dot{\gamma}(a) \in T_aM$.

If $X \in \mathcal{X}(M)$ is a vector field and $F: X \rightarrow Y$ is a diffeomorphism of smooth manifolds, associated with an isomorphism $F^*: \mathcal{C}(N) \rightarrow \mathcal{C}(M)$, then $Y = F_*X$ is a Leibniz-satisfying self-map $Y = (F^*)^{-1}XF^*$ of $\mathcal{C}(N)$.

These definitions by their very nature are global, yet in fact the derivative Xf at a point $a \in M$ depends only on the behavior of the function f and the field X . More precisely, if U is an open subset such that $f \equiv 0$ in U , then for any $X \in \mathcal{X}(M)$ $Xf \equiv 0$ in U .

To make a full use of this observation, we should be able to extend functions which are only “partially defined”, $f \in \mathcal{C}(U)$, $U \subsetneq M$, to functions defined on the whole of M . This is done using the cutoff function.

Definition 4.7. A cutoff function at a point $a \in M$ is a C^∞ -function that is identically equal to 1 in some neighborhood of the point a and vanishes identically outside a larger neighborhood.

Existence of the cutoff functions is at the core of the difference between differential and analytic/algebraic geometry. They cannot exist in the category of algebraic or analytic functions.

Using a cutoff function, for any open subset $U \subseteq M$ and a specified compact subset $K \Subset U$ one can extend smoothly any function $f \in \mathcal{C}(U)$ to a function $\tilde{f} \in \mathcal{C}(M)$ in such a way that $f = \tilde{f}$ on K .

4.4. Flows on manifolds. If $X \in \mathcal{X}(M)$, then under certain assumptions one can construct the flow of X , the family $\{F_X^t\}$ of C^∞ -smooth self-maps of M such that the tangent vector to each curve $\gamma_a: t \mapsto F_X^t(a)$ coincides with $X(a) \in T_aM$.

Theorem 4.1. *If M is compact, then any $X \in \mathcal{X}(M)$ admits the flow defined on $\mathbb{R} \times M \rightarrow M$, i.e., for all times and all initial conditions.*

Proof. This theorem follows from the local theorem of existence and uniqueness for solutions of ODE’s. Each point $a \in U$ admits a neighborhood in which F_X^t is defined in a suitable chart for all sufficiently

small $|t| < \tau_a$, $\tau_a > 0$. Using compactness, one can choose a finite subcovering from $\{U_a\}_{a \in M}$, obtain the minimal guaranteed existence time $\tau = \min_a \tau_a$ and then construct F_X^t for all t by no more than t/τ compositions. \square

Remark 4.1. If M is non-compact, then the flow may be ill-defined for certain combinations (t, a) and some $X \in \mathcal{X}(M)$. However, for any open bounded subset M' (whose closure $\overline{M'} \subsetneq M$ is compact), one can modify X so that it coincides with X on M' and has compact support. Just multiply by a suitable cutoff function.

We already know that diffeomorphisms carry with them (forward or backward) various objects, e.g., curves, functions, vector fields etc. Having a one-parameter family of self-diffeos allows to construct curves in $\mathcal{C}(M)$, $\mathcal{X}(M)$ etc., and the group property provides ways to compute the derivatives with respect to this parameter.

Example 4.1. If $f \in \mathcal{C}(M)$ and $\{F_X^t\}$ the flow of a vector field $X \in \mathcal{X}(M)$, then one can produce a one-parametric family of functions $\{f_t = (F_X^t)^* f\} \subset \mathcal{C}(M)$.

By definition, its derivative, evaluated at a point $a \in M$,

$$g(a) = \left. \frac{d}{dt} \right|_{t=0} f_t(a) = \lim_{t \rightarrow 0} \frac{f(F_X^t(a)) - f(a)}{t} = (Xf)(a)$$

is the Lie derivative of f along X .

A much more interesting example occurs when we try to differentiate one vector field along another.

4.5. Lie derivative of vector fields. Let $X, Y \in \mathcal{X}(M)$ be two vector fields on the same manifold. Let F_X^t be the flow of X . Being a self-diffeo, each F_X^t pushes forward Y to a vector field $Y_t = (F_X^t)_* Y \in \mathcal{X}(M)$. We can therefore attempt to compute the derivative $\left. \frac{d}{dt} \right|_{t=0} Y_t$. However, it is easier to compute the derivative $\left. \frac{d}{dt} \right|_{t=0} (F_X^t)^* Y$, where the pullback operator $(F_X^t)^*$ is defined using the general pullback construction for $\Phi = F_X^t$. Recall that if $\Phi: M \rightarrow N$ is a diffeomorphism (or, more generally, a covering map with nondegenerated differential at each point), then it pulls back vector fields as follows,

$$\Phi^*: \mathcal{X}(N) \rightarrow \mathcal{X}(M), \quad (\Phi^* Y)(x) = (d\Phi(x))^{-1} \cdot Y(\Phi(x)). \quad (3)$$

As always, $d\Phi(x)$ stands for the linear map (differential) from $df(x): T_x M \rightarrow T_{\Phi(x)} N$, represented by the Jacobian matrix $\left(\frac{\partial \Phi}{\partial x}\right)(x)$ in the local coordinates. We will work in the chart near the origin $x = 0$.

Definition 4.8. The Lie derivative along a vector field $X \in \mathcal{X}(M)$ is the operator

$$L_X: \mathcal{X}(M) \rightarrow \mathcal{X}(M), \quad L_X Y = \left. \frac{d}{dt} \right|_{t=0} (F_X^t)^* Y.$$

It is obvious that L_Y is an \mathbb{R} -linear operator on $\mathcal{X}(M)$. A simple argument involving the identity $\Phi^*(fY) = \Phi^* f \Phi^* Y$ for any $\Phi: M \rightarrow N$, $f \in \mathcal{C}(N)$, $Y \in \mathcal{X}(M)$, proves that L_X satisfies the Leibniz rule,

$$L_X(fY) = L_X f \cdot Y + f \cdot L_X Y.$$

Remark 4.2. Strictly speaking, one should use different notation for the pullback operator Φ^* depending whether it is applied to functions or to vector fields. The above identity actually means that one can safely use the same expression in both cases. Moreover, the definition for the Lie derivative L_X above can be considered as “universal”, that is, applicable to all objects that can be naturally pulled back by smooth maps of manifolds. Thus for functions we have

$$L_X f = Xf$$

tautologically.

Theorem 4.2. For any two vector fields $X, Y \in \mathcal{X}(M)$,

$$L_X Y = [X, Y].$$

Proof. We will use the fact that if t, x are sufficiently small, $t \in (\mathbb{R}^1, 0)$, $x \in (\mathbb{R}^n, 0)$,

$$F_X^t(x) = x + tX(x) + (\text{terms small with respect to } t \text{ and } x),$$

so that $dF_X^t(0) = E + tJ_X(0) + o(|t|)$, $J_X = J_X(x) = \left(\frac{\partial X}{\partial x}\right)$ being the Jacobian matrix of X , and, respectively, its inverse is $E - tJ_X(0) + o(|t|)$.

On the other hand, by the Taylor formula

$$Y(F_X^t(x)) = Y(x) + tJ_Y(x) \cdot Y(x) + o(|t| + |x|).$$

Collecting everything together, we see that

$$(L_X Y)(0) = -J_X(0)Y(0) + J_Y(0)X(0),$$

that is, in the local coordinates

$$L_X Y = -\left(\frac{\partial X}{\partial x}\right)Y + \left(\frac{\partial Y}{\partial x}\right)X.$$

If we compare this with the formula for the commutator of two vector fields,

$$[X, Y] = XY - YX = \left(\frac{\partial Y}{\partial x}\right)X - \left(\frac{\partial X}{\partial x}\right)Y,$$

the assertion of the theorem follows. \square

Problem 4.2. Prove this result directly by checking how the field $L_X Y$ acts on an arbitrary test function f by definition.

Using the antisymmetry of the Lie bracket and the Leibniz property proved above, we have an instant corollary.

Corollary 4.1.

$$L_{fX} Y = f L_X Y + (L_X f) Y, \quad \text{i.e.,} \quad L_{fX} = f \cdot L_X + (L_X f) \cdot (\cdot),$$

where the second term is interpreted as the multiplication by the function $L_X f$. \square

A die-hard computator would say that dependence of L_X on X is not tensorial! In particular, the Lie derivative L_X at a point a , considered on vector fields, depends not just on the vector $X(a)$, as was the case with derivation of functions, but on the derivatives of the coordinates of X at a .

4.6. Properties of the Lie derivative. If $L_X f = 0$, then the function f is constant along all trajectories of the vector field. Another way to express this fact is to say that the flow F_X^t preserves f , $(F_X^t)^* f = f$ for all $t \in \mathbb{R}$.

Theorem 4.3. If $L_X Y = 0$, then $(F_X^t)^* Y = Y$ for all $t \in \mathbb{R}$.

Corollary 4.2. If $[X, Y] = 0$, then $F_X^t \circ F_Y^s = F_Y^s \circ F_X^t$ for all $t, s \in \mathbb{R}$.

The identity $[X, Y] = 0$ is a differential identity between components of the vector fields X, Y which is local, i.e., can be independently verified in an arbitrary chart. The commutation of the flows is a “global” fact valid for all times (t, s) . It will be important in integrating distributions, a multidimensional generalization of vector fields.

For the Lie derivatives of functions, one has the identity $(XY - YX)f = [X, Y]f$ by definition (one had to prove that the bracket $[X, Y]$ is indeed a derivation). Rewritten in the alternative notation, this fact reads

$$L_X L_Y - L_Y L_X = L_{[X, Y]}, \quad \text{i.e.,} \quad [L_X, L_Y] = L_{[X, Y]}$$

(the left hand side is the commutator of two differential operators on functions, the right hand side the derivation along the Lie bracket).

It is very tempting to assume that

$$[L_X, L_Y] = L_{[X, Y]} \quad \text{on} \quad \mathcal{X}(M). \quad (4)$$

This is indeed the case: using the formula $L_X = [X, \cdot]$ on vector fields, we reduce the above claim to the identity which is easy to prove.

Theorem 4.4. *For any three vector fields $X, Y, Z \in \mathcal{X}(M)$,*

$$[X, [Y, Z]] - [Y, [X, Z]] = [[X, Y], Z].$$

Proof. This identity, rewritten using the anticommutativity of the Lie bracket, takes the form

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

This is the well-known Jacobi identity. Formally it can be proved by applying the corresponding first order differential operators to a test function f and collect different third order derivatives of the form $(XYZ)f$, $(ZXY)f$ etc. with different signs (all six permutations of the letters X, Y, Z appear twice in the $3 \times 2 \times 2 =$ twelve terms and “mysteriously” cancel each other). \square

What about the Leibniz rule for derivations of vector fields? If we believe that the bracket $[X, Y]$ is in any sense a product of two vector fields, then one should expect the identity

$$L_X[Y, Z] = [L_X Y, Z] + [Y, L_X Z] \tag{5}$$

for any three vector fields. Bah! after substituting $L_X = [X, \cdot]$ it turns into

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$$

is the same familiar Jacobi identity in a different disguise.

In short, a proper natural definition (in our case, the pullback plus the derivation) works beyond its initial scope.

Problem 4.3. Compare Φ and Φ^{-1} , then Φ_* and Φ^* . Compare $(F_X^t)^*$, $(F_{-X}^t)^*$, $(F_X^{-t})^*$ and analogous push-forwards $(F_X^t)_*$, \dots

Bottom line: compute the “push-forward derivative” $\frac{d}{dt}|_{t=0}(F_X^t)_* Y$ and compare it with L_X .

5. Differential forms

5.1. Modules over the algebra $\mathcal{C}(M)$. The space of all (smooth) vector fields $\mathcal{X}(M)$ on a smooth manifold M has the natural structure of a *module* over the ring (\mathbb{R} -algebra) $\mathcal{C}(M)$: this means that besides being the linear space over \mathbb{R} , vector fields can be multiplied by smooth function, yielding the application

$$\mathcal{C}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M), \quad (f, X) \mapsto fX.$$

Problem 5.1. Assuming that $f > 0$, show that the orbits of the flows F_{fX}^t and F_X^t coincide for compact manifolds.

In the same way as any (abstract finite-dimensional) linear space S has its dual S^* as the linear space of all linear functionals (covectors) $\xi: S \rightarrow \mathbb{R}$ equipped with the natural linear operations, one can consider the dual object to vector fields as a family of covectors $a \mapsto \xi(a) \in T_a^*M$, smoothly depending on a .

5.2. Differential 1-forms.

Example 5.1. Let $u \in \mathcal{C}(M)$ be a smooth function. Then the correspondence

$$a \mapsto \xi(a) = du(a)$$

defines a linear functional on each space T_aM : its value on a vector $v \in T_aM$ is the directional derivative of u along v . If we choose a local chart near a in which $v = (v_1, \dots, v_n)$, then the functional is

$$(v_1, \dots, v_n) \mapsto \sum_{i=1}^n \frac{\partial u_i}{\partial x_i}(a) v_i.$$

Clearly, the correspondence $a \mapsto \xi(a)$ is smooth.

Thus the differential du of a smooth function $u \in \mathcal{C}(M)$ can be considered as a \mathbb{R} -linear map

$$du: \mathcal{X}(M) \rightarrow \mathcal{C}(M), \quad du(fX) = f \cdot du(X) \quad \forall f \in \mathcal{C}(M). \quad (6)$$

The tangent space T_aM is n -dimensional and generated by the vectors $\frac{\partial}{\partial x_i}$, $i = 1, \dots, n$. The differentials $dx_i(a): T_aM \rightarrow \mathbb{R}$ form the dual basis in T_a^*M : $dx_i(a)(\frac{\partial}{\partial x_j}) = \delta_{ij}$.

The set of all maps satisfying (6) is itself a module over $\mathcal{C}(M)$ that will be denoted by $\Omega^1(M)$: if $\xi \in \Omega^1(M)$, then there is a naturally defined map $f\xi \in \Omega^1$.

Definition 5.1. The $\mathcal{C}(M)$ -module of maps $\mathcal{X}(M) \rightarrow \mathcal{C}(M)$, the dual module to $\mathcal{X}(M)$, is called the module of differential 1-forms.

It is convenient to denote the pairing between $\Omega^1(M)$ and $\mathcal{X}(M)$ by the angle braces,

$$\forall \xi \in \Omega^1(M), X \in \mathcal{X}(M), \quad \langle \xi, X \rangle \in \mathcal{C}(M).$$

By definition, for any collection of functions $u_k, f_k \in \mathcal{C}(M)$ the expression

$$\sum_k f_k du_k \in \Omega^1(M)$$

is a differential 1-form.

Proposition 5.1. In any coordinate chart $x: U \rightarrow \mathbb{R}^n$ each 1-form can be represented as $\sum f_k(x) dx_k$ with the appropriate smooth functions f_k .

5.3. Action of smooth maps on differential forms. If $F: M \rightarrow N$ is a smooth map, then the pullback operator F^* naturally acts on 1-forms as follows,

$$F^*\left(\sum f_k du_k\right) = \sum F^* f_k \cdot dF^* u_k. \quad (7)$$

This definition is self-consistent.

Problem 5.2. If $F: M \rightarrow N$ is a diffeo and ξ, X live on N , then

$$\langle F^* \xi, F^* X \rangle = F^* \langle \xi, X \rangle, \quad F^*(f\xi) = F^* f \cdot F^* \xi. \quad (8)$$

Prove it.

Definition 5.2. The above definition of the pull-back is simple but ideologically wrong. The correct definition should be as follows.

If $\xi \in \Omega^1(N)$ and $a \in M$, then for any vector $v \in T_a M$ the value $\langle F^* \xi(a), v \rangle$ is equal to $\langle \xi(F(a)), dF(a) \cdot v \rangle$, where $dF(a): T_a M \rightarrow T_{F(a)} N$.

This definition makes sense even if F is not a diffeomorphism. The property (7) becomes then an easy lemma.

Remark 5.1. If $A: P \rightarrow Q$ is a linear map between two abstract linear spaces P, Q , then there is a naturally defined dual (conjugate) map $A^*: Q^* \rightarrow P^*$ between the respective dual spaces, which acts in the “opposite” direction by the formula

$$\langle A^* \eta, v \rangle = \langle \eta, Av \rangle \quad \forall \eta \in Q^*, v \in P.$$

Using this duality, one can rewrite $F^* \xi = (dF)^* \xi$, where $(dF)^*$ stands for the dual of dF .

Example 5.2. Write the last formula “pointwise”.

Note that, unlike the pullback of vector fields, this definition *does not require* the map F to be invertible or even a local diffeo. All smooth maps, even between manifolds of different dimensions, yield well-defined pullbacks.

5.4. Integration of 1-forms. Differential 1-forms are born to be integrated along curves. The question why the definite integral of a real function $f: [a, b] \rightarrow \mathbb{R}$ is denoted

$$\int_a^b f(x) dx$$

necessarily involving mysterious dx , puzzled generations of undergraduate students. The “right” answer is that the object that we integrate is not a function, but rather a differential form $f dx$, where $x: [a, b] \rightarrow \mathbb{R}$ is the coordinate function on $[a, b]$.

If $\gamma: [0, 1] \rightarrow M$ is a smooth curve and $\xi \in \Omega^1(M)$ is a 1-form, then its pullback $\gamma^*\xi$ is a differential 1-form on $\mathbb{R}^1 = \mathbb{R}$ which can be uniquely written as $\gamma^*\xi = f(t) dt$, where t is the “canonical” chart on $\mathbb{R}^1 = \mathbb{R}$.

Definition 5.3. Let $\gamma: [0, 1] \rightarrow M$, $t \mapsto x(t)$, is a smooth (parameterized) curve and $\xi \in \Omega^1(M)$ is a 1-form, then the integral of ξ over $\Gamma = \gamma([0, 1])$ is the real number

$$\int_{\Gamma} \xi = \int_0^1 \gamma^*\xi.$$

This definition prompts a number of natural questions that have expected answers.

Proposition 5.2. The integral is independent on the parametrization: if $t = t(s): [0, 1] \rightarrow [0, 1]$ is a smooth monotonously growing self-map of $[0, 1]$ into itself and $\tilde{\gamma}: s \mapsto \gamma(t(s))$ is a reparametrization of Γ , then

$$\int_{[0,1]} \gamma^*\xi = \int_{[0,1]} \tilde{\gamma}^*\xi. \quad \square$$

A slightly more focussed inspection allows to greatly extend this result for *any* smooth map $\phi: [0, 1] \rightarrow [0, 1]$ which takes the boundary $\{0, 1\}$ into itself. If $\phi(0) = 0$, $\phi(1) = 1$, then such map is called *orientation-preserving* and the corresponding reparametrization does not change the value of the integral. If $\phi(1) = 0$, $\phi(0) = 1$, then the map is called *orientation-reversing*, and the sign of the integral should be changed. To take these nuances into account, we say about *non-parameterized but oriented curves*: these are images $\Gamma = \gamma([0, 1])$ of smooth maps $\gamma: [0, 1] \rightarrow M$ with the marked “start point” $\gamma(0)$ and the “end point” $\gamma(1)$.

Example 5.3. Let $\xi = df$ for $f \in \mathcal{C}(M)$. Then for any oriented curve Γ the integral of ξ can be instantly computed:

$$\int_{\Gamma} df = f(\text{end point}) - f(\text{start point}).$$

This is known to your undergraduate brethren as the Newton-Leibniz theorem. In particular, if the curve Γ is closed, $\gamma(0) = \gamma(1)$, then $\int_{\Gamma} df = 0$.

Remark 5.2. Assume Γ is a smooth curve in a domain $U \subseteq \mathbb{R}^n$. Then Γ can be arbitrarily accurately approximated by a polyline (piecewise-linear curve) by selecting points $A_1, \dots, A_N \in \Gamma$ at very small distances from each other. Then the vectors $v_i = A_{i+1} - A_i \in \mathbb{R}^n$ will have very small norm so that $\langle \xi(A_i), v_i \rangle$ will be very small values, yet the Riemann-like integral sum

$$\sum_{i=1}^{N-1} \langle \xi(A_i), A_{i+1} - A_i \rangle$$

will converge to the limit which, of course, coincides with the integral $\int_{\Gamma} \xi$. The advantage of this approach is the obvious independence of the integral $\int_{\Gamma} \xi$ on the specific parametrization $\gamma(\cdot)$ of Γ .

An alternative way to prove this independence is to refer to the theorem on change of variables in the Riemann integral.

5.5. Lie derivative. If X is a vector field, then the flow F_X^t yields the family of pullbacks $(F_X^t)^*$ on $\Omega^1(M)$ and hence the Lie derivative $L_X: \Omega^1(M) \rightarrow \Omega^1(M)$.

Lemma 5.1.

$$\begin{aligned} \langle L_X \xi, Y \rangle &= L_X \langle \xi, Y \rangle - \langle \xi, L_X Y \rangle = X \langle \xi, Y \rangle - \langle \xi, [X, Y] \rangle, \\ L_X(f\xi) &= fL_X \xi + (L_X f) \cdot \xi. \end{aligned}$$

Proof. This follows from the identity (8) applied to $F = F_X^t$ as $t \rightarrow 0$. □

Example 5.4. What about $L_{fX}\xi$?

Recall that $L_{fX}Y = [fX, Y] = f[X, Y] + (Yf) \cdot X = fL_X Y + \langle df, Y \rangle X$, that is, on the level of operators of $\mathcal{X}(M)$ into itself,

$$L_{fX} = fL_X + \langle df, \cdot \rangle X,$$

thus the application $X \mapsto L_X$ on $\mathcal{X}(M)$ is a non-tensor.

Using this observation, we compute

$$\begin{aligned} \langle L_{fX}\xi, Y \rangle &= L_{fX} \langle \xi, Y \rangle - \langle \xi, L_{fX}Y \rangle \\ &= fL_X \langle \xi, Y \rangle - f \langle \xi, L_X Y \rangle - \langle df, Y \rangle \langle \xi, X \rangle. \end{aligned} \quad (9)$$

Again a non-tensor! $L_{fX} \neq fL_X$ on $\Omega^1(M)$ as well.

The long way to a nice formula goes through a systematic theory involving both (linear) algebraic part and some calculus as well.

6. Multilinear algebra

If $T \simeq \mathbb{R}^n$ is an (abstract) linear space and T^* is its dual, the space of linear functionals on T , then one can make a step further and consider the space of multilinear forms ω , maps

$$T \times \cdots \times T \rightarrow \mathbb{R}, \quad (v_1, \dots, v_k) \mapsto \omega(v_1, \dots, v_k),$$

\mathbb{R} -linear in each argument v_i separately. This space is called the tensor product $T^* \otimes \cdots \otimes T^*$ and in turn is an \mathbb{R} -linear space. If $\xi_1, \dots, \xi_k \in T^*$ are covectors, then the *tensor product*

$$\xi_1 \otimes \cdots \otimes \xi_k: (v_1, \dots, v_k) \mapsto \xi_1(v_1) \cdots \xi_k(v_k)$$

is such a k -linear map, and in general all k -linear maps are sums of such tensor monomials. Note that the tensor product is non-commutative.

A k -valent tensor can be multiplied by an r -valent tensor, using the distributivity law and non-commutative multiplication of monomials, can be multiplied to yield a $k + r$ -valent tensor. This *tensor multiplication* is quite important.

6.1. Antisymmetric multilinear maps. A k -linear map $\omega: T^n \rightarrow \mathbb{R}$ is called symmetric (resp., antisymmetric), if any transposition of its arguments preserves the value (resp., changes only the sign) of the map.

Example 6.1. Let $\delta: (\mathbb{R}^n)^n \rightarrow \mathbb{R}$ be the map which takes vectors v_1, \dots, v_n , expands each vector as a column matrix of its components, and yields the value of the determinant of the resulting $n \times n$ -matrix. Multilinearity and antisymmetry are the basic properties of δ no matter how it was introduced in your favorite undergraduate course.

Quadratic forms are the most known source of symmetric bilinear maps: actually, they are defined using symmetrization of bilinear maps. If $b: T^2 \rightarrow \mathbb{R}$ is any bilinear map, then the map defined by the formula $b^\dagger(v, w) = \frac{1}{2}(b(v, w) + b(w, v))$ is symmetric. Conversely, if $q: T \rightarrow \mathbb{R}$ is a quadratic function, then its polarization $b(v, w) = \frac{1}{2}(q(v + w) - q(v) - q(w))$ is a symmetric bilinear form. In general, the symmetrization of a k -form $b: T^k \rightarrow \mathbb{R}$ is obtained by the averaging over all permutations (\mathfrak{S}_k stands for the group of all permutations on k symbols) of its arguments:

$$b^\dagger(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} b \circ \sigma(v_1, \dots, v_k).$$

Alternatively, a k -linear form b can be subjected to the antisymmetrization (alternation),

$$b^\dagger(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} (-1)^{|\sigma|} b \circ \sigma(v_1, \dots, v_k),$$

where $|\sigma| = \pm 1$ is the parity of the permutation. In the simplest case

$$b^\dagger(v, w) = \frac{1}{2}(b(v, w) - b(w, v)) = -b^\dagger(w, v).$$

For our purposes the antisymmetric case will be more important because of the regular appearance of the determinant. We will denote the set of antisymmetric multilinear forms on T as $\bigwedge^k T^*$.

Example 6.2. Let $\delta_1, \dots, \delta_k$ are k linear independent 1-forms on \mathbb{R}^k , e.g., $\langle \delta_k, v \rangle = k$ -th component of v , and $\delta(v_1, \dots, v_k) = \langle \delta_1, v_1 \rangle \cdots \langle \delta_k, v_k \rangle$ is their tensor product. Then $\delta^\dagger(v_1, \dots, v_k)$ is an antisymmetric k -form which coincides with the determinant up to a constant (equal to one in the e.g.).

This operation turns the union of all k -forms for all $0 \leq k \leq n = \dim T = \dim T^*$ into an *graded exterior algebra*. We list the properties of this algebra: they can all be established by tedious but simple computation.

Definition 6.1. If $\alpha \in \bigwedge^k T^*$ and $\beta \in \bigwedge^l T^*$, then the exterior product (or wedge product) $\alpha \wedge \beta \in \bigwedge^{k+l} T^*$ which is defined by the “antisymmetrization” of the tensor product $\alpha \otimes \beta$, that is,

$$(\alpha \wedge \beta)(v) = \sum_{\sigma \in \mathfrak{S}_{k+l}} (-1)^{|\sigma|} \alpha(\sigma^<(v)) \cdot \beta(\sigma^>(v)) \quad (10)$$

where $\sigma^<$ is the tuple of the first k entries of the permutation σ on $k+l$ symbols $v = (v_1, \dots, v_{k+l})$, and $\sigma^>$ are the last l entries.

Remark 6.1. In some sources the definition of the wedge product (10) involves a coefficient $\frac{1}{k!l!}$ which originates in the alternation of the tensor product. Fortunately, this does not affect the following properties.

Theorem 6.1. *The wedge product possesses the following properties.*

- (1) *It is associative:* $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$.
- (2) *It is \mathbb{R} -distributive:* $\alpha \wedge (\beta + \gamma) = \alpha \wedge \beta + \alpha \wedge \gamma$, $\alpha \wedge (c\beta) = c\alpha \wedge \beta$ for $c \in \mathbb{R}$.
- (3) *It is anticommutative:* for any two 1-forms $\alpha, \beta \in T^* = \bigwedge^1 T^*$, $\alpha \wedge \beta = -\beta \wedge \alpha$. More generally, if $\alpha \in \bigwedge^k T^*$, $\beta \in \bigwedge^l T^*$, then $\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha$.

(4) If $\alpha_1, \dots, \alpha_k \in T^* = \bigwedge^1 T^*$, then

$$(\alpha_1 \wedge \dots \wedge \alpha_k)(v_1, \dots, v_k) = \det \begin{pmatrix} \langle \alpha_1, v_1 \rangle & \dots & \langle \alpha_1, v_k \rangle \\ \vdots & \ddots & \vdots \\ \langle \alpha_k, v_1 \rangle & \dots & \langle \alpha_k, v_k \rangle \end{pmatrix}$$

(5) If $\xi_1, \dots, \xi_n \in T^*$ form a basis in T^* , then any k -form can be uniquely represented as sum of monomials

$$\omega = \sum_{i_1 < \dots < i_k} c_{i_1, \dots, i_k} \xi_{i_1} \wedge \dots \wedge \xi_{i_k},$$

where summation is extended over all monotonously increasing k -subsets $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$.

(6) Hence $\dim \bigwedge^k T^* = \binom{n}{k}$, in particular, $\dim \bigwedge^n T^* = 1$ and any n -form on T is proportional to the “determinant” $\xi_1 \wedge \dots \wedge \xi_n$, called the volume form Υ .

(7) All spaces $\bigwedge^k T^*$ for $k > n$ are trivial (zero-dimensional).

(8) If $A: T \rightarrow T$ is a linear map, then it naturally defines the “adjoint” (dual) linear transformation $A^*: \bigwedge^k T^* \rightarrow \bigwedge^k T^*$ on k -forms by the formula $(A^*\omega)(v_1, \dots, v_k) = \omega(Av_1, \dots, Av_k)$. In particular, $A^*\Upsilon = (\det A) \cdot \Upsilon$.

6.2. Differential k -forms. A differential k -form is a smooth map ω which associates a k -form $\omega(a) \in \bigwedge^k T_a^*(M)$ smoothly depending on a . Algebraically, such form is an antisymmetric module homomorphism of the $\mathcal{C}(M)$ -module $\mathcal{X}(M) \times \dots \times \mathcal{X}(M)$ to $\mathcal{C}(M)$:

$$\omega: (X_1, \dots, X_k) \mapsto \omega(X_1, \dots, X_k) \in \mathcal{C}(M).$$

Sometimes we will use the “duality” notation and denote the right hand side by $\langle \omega, X_1, \dots, X_k \rangle$ which is antisymmetric in the last k arguments.

All linear algebraic constructions can be extended (using the same symbolism) on differential k -forms, in particular, the wedge product.

If $x = (x_1, \dots, x_n)$ is a local coordinate system in U , then the differentials $dx_i \in T_a^*(U)$ are 1-forms linear independent in each cotangent space. It follows then that any k -form can be written in these coordinates as

$$\omega = \sum_{i_1 < \dots < i_k} a_{i_1, \dots, i_k}(x) dx_{i_1} \wedge \dots \wedge dx_{i_k} \quad (11)$$

with $\binom{n}{k}$ smooth coefficients $a_{i_1, \dots, i_k}(x) \in \mathcal{C}(U)$.

We will denote by $\Omega^k(M)$ the $\mathcal{C}(M)$ -module of differential k -forms on a manifold M . The wedge product yields the exterior multiplication $\Omega^k(M) \times \Omega^l(M) \rightarrow \Omega^{k+l}(M)$. The full graded exterior algebra will be denoted by $\Omega^\bullet(M)$.

If $F: M \rightarrow N$ is a smooth map between two manifolds, then it naturally allows to pull back any k -form from N to M by the formula

$$\langle F^*\omega, X_1, \dots, X_k \rangle (a) = \langle \omega, dF(a)X_1, \dots, dF_k(a)X_k \rangle (F(a)).$$

A trivial verification shows that for any monomial k -form $\omega = g df_1 \wedge \dots \wedge df_k$, where $g, f_1, \dots, f_k \in \mathcal{C}(M)$, we have

$$F^*\omega = F^*g \cdot d(F^*f_1) \wedge \dots \wedge d(F^*f_k)$$

(in the right hand side F^* means the pullback $F^*: \mathcal{C}(N) \rightarrow \mathcal{C}(M)$ on smooth functions). More generally, for any two forms on N we have

$$\forall \alpha, \beta \in \Omega^\bullet(M) \quad F^*(\alpha \wedge \beta) = (F^*\alpha) \wedge (F^*\beta).$$

Needless to say, as any product, the wedge product is distributive and derivation of the product yields the Leibniz rule,

$$L_X(\alpha \wedge \beta) = (L_X\alpha) \wedge \beta + \alpha \wedge (L_X\beta).$$

The Lie derivative $L_X\omega$ be computed by applying L_X to the “product” $\langle \omega, Y_1, \dots, Y_k \rangle$ (evaluation of ω on an arbitrary collection of arguments) also using the Leibniz rule,

$$L_X \langle \omega, Y_1, \dots, Y_k \rangle = \langle L_X\omega, Y_1, \dots, Y_k \rangle + \sum_{i=1}^k \langle \omega, Y_1, \dots, L_X Y_i, \dots, Y_k \rangle.$$

Since $L_X Y_i = [X, Y_i]$, the derivative $L_X\omega$ is defined by this identity on any collection of arguments $Y_1, \dots, Y_k \in \mathcal{X}(M)$ in a unique way.

6.3. Recap. Differential k -forms are antisymmetric maps of the $\mathcal{C}(M)$ -module $\mathcal{X}(M) \times \dots \times \mathcal{X}(M)$ into $\mathcal{C}(M)$. The $\mathcal{C}(M)$ -module $\Omega^k(M)$ is generated by monomials $f dg_1 \wedge \dots \wedge dg_k$, $f, g_i \in \mathcal{C}(M)$, where $d: \mathcal{C}(M) \simeq \Omega^0(M) \rightarrow \Omega^1(M)$ is the usual differential as it was introduced for smooth functions. The evaluation is given by the $k \times k$ -determinant.

There is the (algebraic) wedge product defined on the monomials by applying the associative law:

$$(dg_1 \wedge \dots \wedge dg_k) \wedge (df_1 \wedge \dots \wedge df_l) = dg_1 \wedge \dots \wedge df_l.$$

It is associative (by construction) and anticommutative,

$$\omega \wedge \eta = (-1)^{\deg \omega \cdot \deg \eta} \eta \wedge \omega.$$

Besides, it is naturally carried out by the pullback F^* for any smooth map $F: N \rightarrow M$, and any vector field $X \in \mathcal{X}(M)$ induces the Lie derivation(s) $L_X: \Omega^k(M) \rightarrow \Omega^k(M)$ which satisfy the Leibniz rule

$$L_X(\omega \wedge \eta) = (L_X\omega) \wedge \eta + \omega \wedge (L_X\eta),$$

which coincides with the canonical action

$$L_X f = \langle df, X \rangle = Xf \quad \forall f \in \mathcal{C}(M) \simeq \Omega^0(M).$$

7. Exterior differential

If $\omega \in \Omega^k(M)$ is a differential k -form on a smooth manifold M^n (for convenience we identify $\Omega^0(M)$ with $\mathcal{C}(M)$). The Lie derivative $L_X: \Omega^k \rightarrow \Omega^k$ is a nice operator on forms of the same rank k , but the result depends on an extra vector field X . Thus for any form ω we have a map

$$L\omega: (X, Y_1, \dots, Y_k) \mapsto (L_X\omega)(Y_1, \dots, Y_k).$$

It is immediate from the definitions that $L\omega$ is multilinear antisymmetric with respect to the last k arguments. However, with respect to the first argument it is “not linear”¹ (more precisely, it is \mathbb{R} -linear but does not respect the module structure, $L_{fX}\omega \neq fL_X\omega$ for $f \in \mathcal{C}(M)$). This in turn follows from the fact that the Lie derivative $X, Y \mapsto L_X Y = [X, Y]$ is “not linear” in the first argument: $[fX, Y] = f[X, Y] - (L_Y f)X \neq f[X, Y]$. The only exception occurs for $k = 0$:

$$L_{fX}g = f(L_X g) = \langle dg, fX \rangle = f \langle dg, X \rangle.$$

However, one can try and cook from Lie derivatives and a combination of differentials a winning combination that will be linear in X as well.

7.1. Differential of 1-forms. Let $\omega \in \Omega^1$. Then by definition of the Lie derivative,

$$\langle L_X\omega, Y \rangle = \langle L_X\omega, Y \rangle = L_X \langle \omega, Y \rangle - \langle \omega, [X, Y] \rangle.$$

Substituting fX instead of X , we have

$$\begin{aligned} \langle L_{fX}\omega, Y \rangle &= fL_X \langle \omega, Y \rangle - \langle \omega, [fX, Y] \rangle \\ &= fL_X \langle \omega, Y \rangle - f \langle \omega, [X, Y] \rangle + \langle \omega, X \rangle \langle df, Y \rangle. \end{aligned} \quad (12)$$

If we consider the exact 1-form $d\langle \omega, X \rangle$, its dependence on X follows from the Leibniz rule,

$$d\langle \omega, fX \rangle = d(f \langle \omega, X \rangle) = f d\langle \omega, X \rangle + \langle \omega, X \rangle df.$$

¹Not a tensor in the classical language.

The “non-tensorial” part is the same as above, so it disappears if we consider the combination

$$\begin{aligned} \eta_X &= L_X\omega - d\langle\omega, X\rangle \in \Omega^1(M), \\ \eta(X, Y) &= \eta_X(Y) = \langle L_X\omega, Y\rangle - \langle d\langle\omega, X\rangle, Y\rangle. \end{aligned} \tag{13}$$

then η will be linear in X , hence bilinear in X, Y : $\eta_{fX} = f \cdot \eta_X$.

A simple computation shows that

$$\eta(X, Y) = L_X\langle\omega, Y\rangle - L_Y\langle\omega, X\rangle - \langle\omega, [X, Y]\rangle, \tag{14}$$

which instantly implies that $\eta(X, Y) = -\eta(Y, X)$, i.e., our efforts brought the result.

Proposition 7.1. The expression $\eta \in \Omega^2(M)$ is an exterior 2-form: it is bilinear and antisymmetric.

Thus we managed to construct an object that in a sense contains all Lie derivatives of the 1-form ω along all vector fields X , in the same way as the differential of a 0-form df contains all Lie (directional) derivatives L_Xf , and the result $\eta(a): T_a^*M \times T_a^*M \rightarrow \mathbb{R}$ at each point $a \in M$ depends only on the direction (vector) $X(a) \in T_aM$.

Definition 7.1. The 2-form η will be called the *differential* of the form ω . We will provisionally denote the corresponding operator by

$$d': \Omega^1(M) \rightarrow \Omega^2(M)$$

to avoid confusion with $d: \Omega^0(M) \rightarrow \Omega^1(M)$.

Theorem 7.1. *The operator d' possesses the following algebraic properties:*

- (1) *The dependence of $d'\omega$ on ω is \mathbb{R} -linear.*
- (2) *The Leibniz rule: $d'(f\omega) = df \wedge \omega + f d'\omega$.*
- (3) *The composition $d'd: \Omega^0 \rightarrow \Omega^2$ is identically zero.*

Proof. The first claim is obvious. The Leibniz rule follows from the (14) and the Leibniz rule for L_X, L_Y : the terms which are not divisible by f are $L_Xf \cdot \langle\omega, Y\rangle - L_Yf \langle\omega, X\rangle = \langle df, X\rangle \langle\omega, Y\rangle - \langle df, Y\rangle \langle\omega, X\rangle = (df \wedge \omega)(X, Y)$.

The last assertion, somewhat unexpected, follows from the representation (14).

$$d'df = L_XL_Yf - L_YL_Xf - L_{[X, Y]}f = 0$$

by the mere definition of the commutator $[X, Y]$. □

7.2. Antiderivation. A convenient way to write (and understand) such identities is via the *operator of evaluation* or, more generally, an operator of insertion (also known as contraction). By definition, if $X \in \mathcal{X}$, then

$$i_X: \Omega^1 \rightarrow \Omega^0, \quad i_X \omega = \langle \omega, X \rangle,$$

and similarly,

$$i'_X: \Omega^2 \rightarrow \Omega^1, \quad (i'_X) \eta = \eta(X, \bullet) = \langle \eta, X, \bullet \rangle.$$

In these terms the definition of the differential d implies the identity

$$L_X = i_X d \quad \text{as operator } \Omega^0 \rightarrow \Omega^1 \rightarrow \Omega^0.$$

The representation (13) can be rewritten now in terms of the *four* operators,

$$L_X = i'_X d + d' i_X \quad \text{as operator } \Omega^1 \rightarrow \Omega^2 \rightarrow \Omega^1. \quad (15)$$

Actually, one can instantly generalize the operator of contraction i_X to act on any k -forms (except for $k = 0$).

Definition 7.2.

$$i_X: \Omega^k(M) \rightarrow \Omega^{k-1}(M), \quad (i_X \omega)(Y_1, \dots, Y_{k-1}) = \omega(X, Y_1, \dots, Y_{k-1}).$$

The properties of this operator are obvious.

Theorem 7.2 (Leibniz rule for the antiderivation).

$$i_X(\omega \wedge \eta) = (i_X \omega) \wedge \eta + (-1)^{\deg \omega} \omega \wedge (i_X \eta), \quad (16)$$

besides, $i_X i_Y = -i_Y i_X$.

Of course, in addition to the operators i_X decreasing the degree, one has dual operators which increase the degree. If $\xi \in \Omega^1$,

$$\xi^\wedge: \Omega^k \rightarrow \Omega^{k+1}, \quad \omega \longmapsto \xi \wedge \omega. \quad (17)$$

Problem 7.1. Find linear algebraic identities relating operators ξ^\wedge and i_X .

8. Playing good old doctor Grassmann...

Can we generalize the operators d, d' to “derivatives” acting on higher order forms and increasing their degree by 1? In other words, we look for maps

$$\Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d'} \Omega^2(M) \xrightarrow{d''} \Omega^3(M) \rightarrow \dots \rightarrow \Omega^n(M) \quad (18)$$

satisfying the conditions of Theorem 7.1?

The answer is positive, and such generalization can be constructed by two equivalent ways.

First, we can postulate the Leibniz rule and require that

$$d(\omega \wedge \xi) = (d\omega) \wedge \xi + (-1)^{\deg \omega} \omega \wedge (d\xi) \quad \text{for any two forms } \omega, \xi. \quad (19)$$

Theorem 8.1. *There exists a unique family of \mathbb{R} -linear operators*

$$d = d^{(k)} : \Omega^k(M) \rightarrow \Omega^{k+1}(M)$$

such that:

- (1) $d^{(0)}$ coincides with the differential on $\Omega^0(M)$ and with the exterior differential $d' = d^{(1)}$ on $\Omega^1(M)$, in particular, $d'df = 0$ for any $f \in \Omega^0(M) \simeq \mathcal{C}(M)$,
- (2) d satisfies the Leibniz rule (19).

Proof. For any k -form

$$\omega = \sum_{i_1 < \dots < i_k} a_{i_1, \dots, i_k}(x) dx_{i_1} \wedge \dots \wedge dx_{i_k} \in \Omega^k(M)$$

define

$$d\omega = \sum_{i_1 < \dots < i_k} da_{i_1, \dots, i_k} \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} \in \Omega^{k+1}(M).$$

One can then expand each differential

$$da_{i_1, \dots, i_k} = \sum_{j=1}^n \frac{\partial a_{i_1, \dots, i_k}}{\partial x_j} dx_j,$$

keep only j different from i_1, \dots, i_k , and then use the anticommutative law to place each term $dx_j \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}$ to the non-decreasing order. \square

We will naturally simplify the notation and write d instead of $d^{(k)}$ everywhere: the theorem asserts that we don't run into any trouble.

Example 8.1. For $f \in \Omega_0$ we have $df = \sum_i \frac{\partial f}{\partial x_i} dx_i$ and

$$d^2 f = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i \wedge dx_j = 0,$$

since the terms (i, j) and (j, i) cancel each other.

Theorem 8.2 (The Cartan “magic” formula, a.k.a. the homotopy formula). *The formula (15) holds for forms of any degree,*

$$L_X = di_X + i_X d. \quad (20)$$

Proof. The formula is true for 1-forms, and for any monomial k -form $\omega = f dg_1 \wedge \cdots \wedge dg_k$ one can check it directly assuming that $d\omega = df \wedge dg_1 \wedge \cdots \wedge dg_k$ and (repeatedly) the Leibniz rule for i_X and L_X . \square

The alternative can be to *postulate* the Cartan formula for forms of arbitrary degree. Then, rewriting it as

$$i_X d\omega = L_X \omega - di_X \omega,$$

we can *define* $d\omega$ for any k -form ω and any vector field X . A simple check similar to that made at the beginning of this section shows that the expression for $d\omega$ defined by this way, will be indeed a multilinear antisymmetric $(k+1)$ -form.

One way or another, we have constructed the (co)chain complex

$$\Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \Omega^2(M) \xrightarrow{d} \Omega^3(M) \rightarrow \cdots \xrightarrow{d} \Omega^n(M) \quad (21)$$

with the property $d^2 = 0$.

One can explicitly express the value of $d\omega$ of k -form on a tuple of $k+1$ vector fields X_0, X_1, \dots, X_k in a symmetric (or rather antisymmetric) way. However, it is always better to use the inductive definition rather than the following explicit double sum.

Theorem 8.3.

$$\begin{aligned} d\omega(X_0, X_1, \dots, X_k) = & \sum_{i=0}^k (-1)^i L_{X_i}(\omega(\dots, \hat{X}_i, \dots)) + \\ & \sum_{i < j} (-1)^{i+j} \omega([X_i, X_j], \dots, \hat{X}_i, \dots, \hat{X}_j, \dots), \end{aligned} \quad (22)$$

where hats mean that the corresponding arguments are struck out from the list (X_0, \dots, X_k) . \square

Remark 8.1. There is a nice mnemonic way to remember such formulas, including the Cartan formula. To each of the operators i, L, d we assign the *degree*, equal to $-1, 0$ and $+1$ respectively, in accordance of how they affect the degrees of the forms on which they act.

Define the *twisted commutator* of two operators D_1, D_2 as follows,

$$\{D_1, D_2\} = D_1 D_2 - (-1)^{\deg D_1 \deg D_2} D_2 D_1.$$

Thus the twisted commutator is sometimes the usual commutator, sometimes the “average” $D_1 D_2 + D_2 D_1$.

Then we have the following table of relations:

$$\begin{aligned} \{i_X, d\} &= L_X, \\ \{i_X, i_Y\} &= 0, \\ \{d, d\} &= 0. \end{aligned} \quad (23)$$

The twisted commutators involving L_X are the same as the usual commutators, since $\deg L_X = 0$:

$$\begin{aligned}\{L_X, L_Y\} &= L_{[X, Y]}, \\ \{L_X, i_Y\} &= i_{[X, Y]}, \\ \{L_X, d\} &= 0.\end{aligned}\tag{24}$$

Problem 8.1. Compute the twisted commutators involving the operators ξ^\wedge , see (17).

References

The best source for this subject is the book [A]. The book [KMS] is more formal, but also very useful.

- [A] V. I. Arnold, *Matematicheskie metody klassicheskoi mekhaniki*, 3rd ed., “Nauka”, Moscow, 1989 (Russian). MR1037020

From the MReview:

Let S_1 ={the most influential books of the second half of this century}, S_2 ={the most frequently quoted books}, S_3 ={books that have the highest probability of surviving into the 21st century}, S_4 ={books that are very useful in teaching}, S_5 ={books characterized by a very strong personal style}, S_6 ={books that provide a delightful reading experience}, A =the book under review. **Proposition.** $A \in \bigcap_{i=1}^6 S_i$.

Needless to say, there exist translations into English, French, Polish, German, . . . , most of them available on the Web.

- [KMS] I. Kolář, P. W. Michor, and J. Slovák, *Natural operations in differential geometry*, Springer-Verlag, Berlin, 1993. MR1202431

9. De Rham cohomology

9.1. Recap. Poincaré lemma for differential forms. Let U be a star-shaped domain in \mathbb{R}^n and $\omega \in \Omega^k(U)$ is a closed form, $d\omega = 0$. We construct explicitly a $(k - 1)$ -form α such that $\omega = d\alpha$.

Consider the *linear homotopy* contracting U to the origin. By definition, this is a one-parametric family of smooth maps $H_t: x \mapsto tx$, $x \in U$, $t \in [0, 1]$. Obviously, $H_1^*\omega = \omega$, $H_0^*\omega = 0$. The derivative $\frac{d}{dt}H_t^*$ can be easily expressed through the Lie derivative along the Euler (radial) vector field $E = \sum_1^n x_i \frac{\partial}{\partial x_i}$, since $H_t = F_E^{\ln t}$ (the r.h.s. is the flow

of E):

$$\frac{d}{dt} H_t^* = \frac{d}{dt} F_E^{\ln t} = \frac{1}{t} (F_E^{\ln t})^* L_E = \frac{1}{t} H_t^* (i_E d + d i_E).$$

Note that despite the “singular” coefficient $\frac{1}{t}$, the result is a smooth form by construction (the incredulous can verify it by direct inspection).

Applying this identity to the closed form ω and integrating the result in t along $[0, 1]$ we see that

$$\omega = H_1^* \omega - H_0^* \omega = \int_0^1 \frac{d}{dt} H_t^* \omega dt = \int_0^1 \frac{1}{t} H_t^* d i_E \omega dt = d \int_0^1 H_t^* i_E \omega dt.$$

Note that the homotopy H_t is exactly the same that was used to construct the cone over a chain σ in the “geometric” proof of the Poincaré lemma.

9.2. De Rham complex. The family of infinite-dimensional spaces equipped with the operator $d^2 = 0$. The cohomology of this complex is by definition $\text{Ker } d / \text{Im } d$, i.e., the set of obstructions preventing a closed form to be exact (integrable).

From the Stokes theorem it is clear that any closed form which has a nonzero integral over a cycle (chain with zero boundary) cannot be integrated. The big “if” is whether this is indeed an obstruction (perhaps, such forms don’t exist at all) and whether this is a unique obstruction to integrability.

The question is settled by the de Rham theorem, but there is much more in stock.

9.3. How to calculate the cohomology. Assume that $M = \bigcup_i U_i$ is an open covering of M such that all U_i and their nonempty pairwise intersections $U_{ij} = U_i \cap U_j$ are diffeomorphic to the unit ball (hence subject to the Poincaré lemma).

Consider a closed 1-form $\omega \in \Omega^1(M)$. As we know, in general ω may be non-exact, but how can one measure the obstruction to non-exactness?

By the Poincaré lemma, all restrictions $\omega|_{U_i}$ are exact: there exists a collection of smooth functions $f_i \in \mathcal{C}(U_i)$, such that $df_i = \omega$ in U_i . Such functions are defined modulo constant terms: $\tilde{f}_i = f_i + c_i$ would do the job for any constants $c_i \in \mathbb{R}$.

On each (nonempty) intersection U_{ij} the differentials of two functions must be the same, $d(f_i - f_j) = 0$. Therefore there exist constants $r_{ij} = -r_{ji} \in \mathbb{R}$ such that $f_i - f_j = r_{ij}$. This holds only if U_{ij} is

connected, but we had explicitly assumed this. The values r_{ij} are essentially determined by the form ω .

Suppose that we can find constants $c_i \in \mathbb{R}$ in such a way that

$$c_i - c_j = r_{ij} = -r_{ji} = c_j - c_i \quad \text{for all } i, j \text{ such that } U_{ij} \neq \emptyset.$$

Then replacing the functions f_i by the functions $\tilde{f}_i = f_i + c_i$, we see that $\tilde{f}_i = \tilde{f}_j$ on U_{ij} , that is, the “local primitives” f_i after correction by suitable constants c_i match each other on the intersections and hence produce the global primitive \tilde{f} such that $d\tilde{f} = \omega$.

Thus the question about “obstruction to integrability” is reduced to the following problem from linear algebra: when the above system of linear algebraic equations solvable? A more accurate (and interesting) form of the same question is as follows.

Let $\{U_i\}_{i \in I_1}$ be the covering as above, and let $I_2 = \{(i, j) : U_{ij} \neq \emptyset\} \subseteq I_1 \times I_1$ be the set of indices corresponding to nonempty pairwise intersections. Consider the linear map

$$\mathbb{R}^{I_1} \rightarrow \mathbb{R}^{I_2}, \quad \{c_i\} \mapsto \{c_i - c_j\}.$$

What is the corank of this map, i.e., the codimension of its image?

Note that the “structure” of the map is extremely simple, the devil is in the combinatorics of indices which is ultimately determined by the covering.

Example 9.1. Consider the circle which is covered by two maps. . .

9.4. What about 2-forms? In this case we will have to consider triple intersections.

Primitives of a closed 2-form $\omega \in \Omega^2(M)$ are 1-forms $\xi_i \in \Omega^1(U_i)$ which are well defined in U_i . As before, they may disagree on the pairwise intersections U_{ij} :

$$\xi_i - \xi_j = \alpha_{ij} \in \Omega^1(U_{ij}).$$

Obviously, $d\alpha_{ij} = 0$. We want to twist the closed 1-forms ξ_i by exact 1-forms $\xi_i + df_i$ so that after the twist we will have $\alpha_{ij} = 0$. To that end, we have to ensure that

$$df_i - df_j = \alpha_{ij} \quad \text{in } U_{ij}.$$

Since $d\alpha_{ij} = 0$, we can find smooth functions g_{ij} such that $df_i - df_j = dg_{ij}$. However, in general this does not imply yet that we can find f_i such that $g_{ij} = f_i - f_j$: if it were the case, we would have $g_{ij} + g_{jk} + g_{ki} = 0$, which is not guaranteed: in general, the above sum is a nontrivial constant r_{ijk} . However, if we can twist g_{ij} by suitable constants $c_{ij} \in \mathbb{R}$ so that

$$c_{ij} + c_{jk} + c_{ki} = r_{ijk}$$

over all triple intersections, then we can resolve the last system of equations and construct the functions g_{ij} and f_i as required.

9.5. Čech cohomology. In the initial approximation the Čech cohomology is defined for acyclical coverings. Those can be thought as coverings of a manifold by open sets $\{U_i\}$ such that all nonempty finite intersections are diffeomorphic to the open ball. Locally this is not a problem: convex bodies satisfy this condition, so a covering that is fine enough can be assumed acyclic without loss of generality.

Definition 9.1. A (real) k -cochain \mathbf{c} is a collection of real numbers $\mathbf{c} = \{c_{i_0 i_1 \dots i_k}\}$ associated with all non-empty intersections $U_{i_0 \dots i_k} = U_{i_0} \cap \dots \cap U_{i_k}$ which is antisymmetric with respect to all permutations of the indices.

For each $(k-1)$ -cochain \mathbf{c} its coboundary is a k -cochain $\delta\mathbf{c}$ defined as follows,

$$(\delta\mathbf{c})_{i_0 \dots i_k} = \sum_{r=0}^k (-1)^r c_{i_0 \dots \widehat{i_r} \dots i_k}$$

(the r th index is omitted).

Proposition 9.1.

$$\delta\delta = 0.$$

10. Riemannian manifolds

Recall that for any smooth manifold M , $\dim M = n$, the union $TM = \bigcup_{a \in M} T_a M$, called the *tangent bundle*, is itself a smooth manifold, $\dim TM = 2n$.

Example 10.1. Prove this and construct an atlas of charts on TM .

Recall that each bilinear form $B: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ can be made into a quadratic form $Q: \mathbb{R}^n \rightarrow \mathbb{R}$ by restricting on the diagonal, $Q(v) = B(v, v)$. Conversely, any quadratic form Q can be *polarized* to a *symmetric* bilinear form $B(u, v) = \frac{1}{2}(Q(u+v) - Q(u) - Q(v))$. It is positive definite, if $Q(v) \geq 0$ and $Q(v) = 0 \iff v = 0$.

10.1. Riemannian manifolds: definitions.

Definition 10.1. A Riemannian metric on a manifold M is a smooth function $g: TM \rightarrow \mathbb{R}_+$ which, when restricted on any tangent space $T_a M$, is a positive definite quadratic form. A Riemannian manifold is a manifold equipped with a specific Riemannian metric.

Example 10.2. The Euclidean space \mathbb{R}^n equipped with the same standard quadratic form $g(v) = \langle v, v \rangle = \sum_{i=1}^n v_i^2$ for any $v = (v_1, \dots, v_n) \in T_a\mathbb{R}^n \simeq \mathbb{R}^n$.

The vast source of examples is provided by the *induced* structures: if M is a Riemannian manifold and $N \subseteq M$ is a submanifold. Then $TN \subseteq TM$, and the restriction of g on TN is a Riemannian metric making N into a Riemannian manifold.

Thus any smooth submanifold of \mathbb{R}^n , say, the unit sphere \mathbb{S}^{n-1} inherits the structure of a Riemannian manifold.

The same manifold can be equipped with different Riemannian metrics. For instance, the torus obtained by rotation of a circle around an axis in \mathbb{R}^3 is embedded in \mathbb{R}^3 in the natural way and so inherits the Riemannian structure.

On the other hand, if we consider the quotient $\mathbb{T}_{a,b}^2 = \mathbb{R}^2/a\mathbb{Z} + b\mathbb{Z}$ for a pair of linear independent vectors $a, b \in \mathbb{R}^2$, then it is naturally equipped by the Riemann metric in \mathbb{R}^2 , since all translations preserve the scalar product. These are called *flat* tori.

Definition 10.2. Two Riemannian manifolds (M, g) and (N, h) are *isometric*, if there exists a diffeomorphism F between them, which transforms one Riemannian metric into another.

If $F: (M, g) \rightarrow (N, h)$ is a differentiable injective² map such $F^*h = g$, then we say that M is isometrically embedded in N .

10.2. What can be done with a Riemannian metric? First, it allows to identify $\mathcal{X}(M)$ with $\Omega^1(M)$: a vector field X and a 1-form ξ are dual, if for any other vector field $Y \in \mathcal{X}(M)$ one has

$$g(X, Y) = \xi(Y).$$

Example 10.3. The vector field $\text{grad } f$ is dual to the 1-form df .

A *huge* generalization of this example is the *Hodge star operator* making $\Omega^k(M)$ and $\Omega^{n-k}(M)$ into dual spaces. Will be discussed later.

Second, we can measure lengths of the smooth curves and angles between them: by definition, the length of a smooth curve $\gamma: [0, 1] \rightarrow M$ is the integral

$$L(\gamma) = \int_0^1 \sqrt{g(\dot{\gamma}(t), \dot{\gamma}(t))} dt > 0.$$

Note that the length *is not* given by the integral of a 1-form, so $L(\gamma) = L(-\gamma)$, but is independent of the parametrization of the curve. A

²Why there cannot be isometric maps which *decrease* the dimension?

similar expression is given by the *action*³, the integral

$$A(\gamma) = \int_0^1 g(\dot{\gamma}(t), \dot{\gamma}(t)) dt = \int_{\gamma} \xi_{\gamma}, \quad \xi_{\gamma} = i_{\dot{\gamma}}g = g(\dot{\gamma}, \cdot) \in \Omega^1(M),$$

where ξ_{γ} is the 1-form dual to the velocity vector field $\dot{\gamma}$ defined along γ . The advantage of action is differentiability of the integrand, disadvantage is the explicit dependence on parametrization.

Having lengths, one can meaningfully talk about shortest curves connecting two given endpoints. Such curves are called geodesic curves. An smooth infinite curve is called geodesic, if it can be split into shortest curves (spell out the formal definition!).

Example 10.4. In the Euclidean space the shortest curves are line segments, and infinite geodesics are straight lines.

One can use simple geometric arguments to show that on the standard round sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ the shortest curves are arcs of the large circles (sections of the sphere by planes through the origin): it is enough to verify that such curves must be planar. The arc of length $> \pi$ is not shortest anymore (its complement is), and large circles (without the endpoints) are geodesics. There are no non-closed geodesics on the sphere!

There are no topological obstructions for a manifold to be Riemannian.

Theorem 10.1. *Any smooth manifold can be equipped with a Riemannian metric.*

Proof. Let $\{U_i\}$ be a locally finite covering of M by charts $x_i: U_i \rightarrow \mathbb{R}^n$ and $1 = \sum_i \psi_i$ the subordinate partition of unity, $\mathcal{C}(M) \ni \psi \geq 0$, $\text{supp } \psi_i \subseteq U_i$. In each chart we have metric g_i pulled back from \mathbb{R}^n . The sum $\sum \psi_i g_i$ yields a Riemannian metric which is nondegenerate.

Note that this argument would fail if instead of the positive definite quadratic form we were looking for a metric with a different signature! \square

10.3. Parallel transport and covariant derivation. There is a rather large group of isometries of the Euclidean space (\mathbb{R}^n, g) , $g(x, v) = \langle v, v \rangle$, containing the subgroup (isomorphic to \mathbb{R}^n) of parallel translations.

The existence of the parallel transport allows one to calculate the derivative of a vector field Y along any smooth curve as follows: if

³Sometimes this integral is referred to as the energy and preceded by the coefficient $\frac{1}{2}$.

the curve is given by the parametrization $\gamma : [0, 1] \rightarrow \mathbb{R}^n$, $t \mapsto x(t) = (x_1(t), \dots, x_n(t))$ and the vector field is given by its coordinate functions $(Y_1(x), \dots, Y_n(x))$, then we can take a composition $Y(\gamma(t)) : [0, 1] \rightarrow \mathbb{R}^n$ and compute the velocity $\frac{d}{d\gamma}Y = \frac{d}{dt}Y(\gamma(t))$. By the chain rule, it can be expressed through the differential operator

$$\bar{\nabla}_X Y = (L_X Y_1, \dots, L_X Y_n), \quad X, Y, \bar{\nabla}_X Y \in \mathcal{X}(\mathbb{R}^n), \quad (25)$$

as the derivation along the velocity vector $\dot{\gamma}$,

$$\frac{d}{d\gamma}Y = \bar{\nabla}_{\dot{\gamma}}Y.$$

Looking at the operator $\bar{\nabla}$, we immediately see that it satisfies a number of properties:

- (1) $\bar{\nabla}$ is \mathbb{R} -linear in both arguments X, Y ,
- (2) $\bar{\nabla}_X$ respects the Leibniz rule: $\bar{\nabla}_X(fY) = (\bar{\nabla}_X f)Y + f(\bar{\nabla}_X Y)$. There is only one way to interpret $\bar{\nabla}_X f$ for $f \in \mathcal{C}(M)$ as $L_X f$.
- (3) $\bar{\nabla}_{fX} Y = f\bar{\nabla}_X Y$, that is, the operator $\bar{\nabla}$ acts on vector fields in the same way as L_X acts on functions (such behavior is called *tensorial*).
- (4) Moreover, if $Y, Z \in \mathcal{X}(\mathbb{R}^n)$ are two vector fields, then the Leibniz rule holds for the scalar product $\langle \cdot, \cdot \rangle$:

$$\bar{\nabla}_X \langle Y, Z \rangle = L_X \langle Y, Z \rangle = \langle \bar{\nabla}_X Y, Z \rangle + \langle Y, \bar{\nabla}_X Z \rangle. \quad (26)$$

By definition, $\bar{\nabla}_X$ acts on vector fields in the same way as $X = L_X$ acts on scalar functions. In particular,

$$\bar{\nabla}_X Y - \bar{\nabla}_Y X = [X, Y]. \quad (27)$$

Definition 10.3. A differential operator $\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M)$ satisfying the properties (1)–(3) above, is called the *covariant derivation*. The image $\nabla_X Y$ is the *covariant derivative* of the vector field Y along the vector field X .

If γ is a (piecewise)-smooth curve, than the “vector field” Y defined only on the image of γ , is called *constant along γ* , if $\nabla_{\dot{\gamma}}Y = 0$. For any such field the value $Y(\gamma(1))$ is called the *parallel transport* of $Y(\gamma(0))$ along γ .

We say that ∇ defines a *connexion* on the manifold M . It is called *symmetric*, if (27) holds. If M is a Riemannian manifold and (26) holds, we say about the Riemannian connexion.

The property (26) means that the parallel transport in \mathbb{R}^n is an isometry between $T_{\gamma(0)}\mathbb{R}^n$ and $T_{\gamma(1)}\mathbb{R}^n$.

10.4. Induced covariant derivation. We will show how the derivation $\bar{\nabla}_X Y$ can be “restricted” to vector fields tangent to a hypersurface $M \subseteq \mathbb{R}^n$, producing a vector field which is also tangent to M .

If $M^{n-1} \subset \mathbb{R}^n$ is a smooth hypersurface, then the parallel transport in \mathbb{R}^n does not map tangent subspaces into themselves. On the infinitesimal level, the covariant derivative $\bar{\nabla}_X Y$ for two vector fields $X, Y \in \mathcal{X}(M)$ tangent to M is not necessarily tangent to M .

The required correction is both natural and minimal.

Definition 10.4. The covariant derivative $\nabla_X Y$ of a vector field $Y \in \mathcal{X}(M)$ along $X \in \mathcal{X}(M)$ is the orthonormal projection of $\bar{\nabla}_X Y \in \mathcal{X}(\mathbb{R}^n)$ onto $TM \subseteq T\mathbb{R}^n \simeq \mathbb{R}^n \times \mathbb{R}^n$. Of course, it depends on the hypersurface M .

A vector field $Y \in \mathcal{X}(M)$ is said to be parallel along a curve $\gamma : [0, 1] \rightarrow M$, if $\nabla_{\dot{\gamma}} Y = 0$. This defines the linear operator between $T_{\gamma(0)}\mathbb{R}^n$ and $T_{\gamma(1)}\mathbb{R}^n$, called the *parallel transport*.

Example 10.5. Let $E_1, \dots, E_n \in \mathcal{X}(M)$ be (locally) linear independent vector fields on M , say, coming from the derivations $\partial_{x_1}, \dots, \partial_{x_n}$ in some chart. Then we can define n^3 functions on M by the conditions,

$$\nabla_{E_i} E_j = \sum_{k=1}^n \Gamma_{ij}^k E_k, \quad (28)$$

expanding the covariant derivatives in the basis $\{E_i\}$. The functions Γ_{ij}^k are called the *Christoffel symbols*. If the connexion ∇ is symmetric, then $\Gamma_{ij}^k = \Gamma_{ji}^k$.

Assume that the curve γ is given by its smooth parametric equation $t \mapsto \gamma(t)$. Then a vector field $Y = \sum y_i(t) E_i(t)$ is parallel along γ , if and only if $\nabla_{\dot{\gamma}} Y = 0$. Expanding this identity using the Leibniz rule, we obtain a system of linear homogeneous ODEs for the functions $y_i(t)$, which can be solved with any initial condition.

The fundamental question is as follows. Embedding of M into \mathbb{R}^n induces the Riemannian metric on M . The parallel transport and covariant derivation are explicitly defined in terms of the embedding (which is much more information). Can one determine the result of the transport in terms of the induced metric only? The answer obtained by Gauss, the celebrated *Theorema Egregium*, is affirmative. We outline the steps towards explaining this result.

Let $M^{n-1} \subset \mathbb{R}^n$ be a smooth hypersurface and N a unit normal vector field on it:

$$\forall a \in M \quad N(a) \in T_a\mathbb{R}^n, \quad \langle N(a), N(a) \rangle = 1, \quad \langle N(a), \cdot \rangle|_{T_a M} = 0.$$

The map $a \mapsto N(a)$ is sometimes called the Gauss map $M \rightarrow \mathbb{S}^{n-1} \subseteq \mathbb{R}^n$.

For any $X \in \mathcal{X}(M)$ the “ambient covariant derivative” $\bar{\nabla}_X N$ is tangent to M . Indeed, by the Leibniz rule

$$\langle \bar{\nabla}_X N, N \rangle = \frac{1}{2} L_X \langle N, N \rangle = 0, \quad \text{hence } N \perp \bar{\nabla}_X N \in \mathcal{X}(M).$$

Definition 10.5. The *Weingarten operator* is the linear operator $W_a: T_a M \rightarrow T_a M$ which sends a vector v into $\bar{\nabla}_v N$ (since $\bar{\nabla}$ is a covariant derivative, one can choose any $X \in \mathcal{X}(M)$ such that $X(a) = v$, and compute $\bar{\nabla}_X N$). For a vector field $X \in \mathcal{X}(M)$ we denote by WX the field $a \mapsto W_a X(a) \in T_a M$.

The Weingarten operator W_a is the Jacobian (derivative) of the Gauss map at a point $a \in M$.

Example 10.6. For the round unit sphere $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ the Gauss map is the identity, so the Weingarten operator is also identical.

Lemma 10.1. *The Weingarten operator is self-adjoint in the induced metric on $T_a M$.*

Proof. Let $X, Y \in \mathcal{X}(M)$ be any two vector fields tangent to M . We want to prove that

$$\langle WX, Y \rangle = \langle X, WY \rangle.$$

Without loss of generality we may assume that X, Y, N are defined in a neighborhood of M . Then, using twice the Leibniz rule (26) and the symmetry condition (27) of $\bar{\nabla}$, we conclude that

$$\begin{aligned} \langle WX, Y \rangle - \langle X, WY \rangle &= \langle \bar{\nabla}_X N, Y \rangle - \langle X, \bar{\nabla}_Y N \rangle \\ &= (\bar{\nabla}_X \langle Y, N \rangle - \langle \bar{\nabla}_X Y, N \rangle) - (\bar{\nabla}_Y \langle X, N \rangle - \langle N, \bar{\nabla}_Y X \rangle) \\ &= 0 - 0 - \langle N, \bar{\nabla}_X Y - \bar{\nabla}_Y X \rangle = -\langle N, [X, Y] \rangle. \end{aligned}$$

But the commutator of two vector fields X, Y tangent to M , is again tangent to M , hence the result is zero. \square

Definition 10.6. The eigenvalues $\lambda_1(a), \dots, \lambda_{n-1}(a)$ of the Weingarten operator L_a are called the *principal curvatures* of the hypersurface. By construction, they are functions of the point on M . While the principal curvatures themselves may be non-smooth, their symmetric combinations (coefficients of the characteristic polynomial $\det(\lambda - L_a)$), are smooth. In particular, smooth are the determinant $K(a) = \prod_1^{n-1} \lambda_i(a)$, called the *Gaussian curvature*, and the trace $H(a) = \sum_1^{n-1} \lambda_i(a)$ called the *mean curvature* of M . The corresponding eigenspaces of $T_p M$ are called *directions of curvature* (they are pairwise orthogonal for different eigenvalues).

Note that all these definitions are still non-intrinsic: they explicitly depend on the embedding of M into \mathbb{R}^n .

The meaning of the curvatures can be seen from the following construction. Let $X_i(a) \in T_aM$ be the eigenvector of $W = W_a$, corresponding to the eigenvalue $\lambda_i(a)$. Consider the 2-plane $\Pi_i(a)$ in \mathbb{R}^n spanned by $N(a)$ and $X_i(a)$. Then the Gauss map restricted on the 1-dimensional subspace $\mathbb{R}X_i(a) \subset T_aM$ and its derivative can be instantly computed in terms of the osculating circle of the section $\Pi_i(a) \cap M$ as the inverse radius of this circle.

10.5. The Gauss equation. Using the Weingarten operator, one can easily express the induced covariant derivative as it was defined in Definition 10.4: to compute the orthogonal projection, one has to add to $\bar{\nabla}_X Y$ the normal vector N with a suitable coefficient,

$$\nabla_X Y = \bar{\nabla}_X Y - \langle \bar{\nabla}_X Y, N \rangle N = \bar{\nabla}_X Y + \langle Y, \bar{\nabla}_X N \rangle N = \bar{\nabla}_X Y + \langle WX, Y \rangle N,$$

since

$$0 = \bar{\nabla}_X \langle Y, N \rangle = \langle \bar{\nabla}_X Y, N \rangle + \langle Y, \bar{\nabla}_X N \rangle = \langle Y, WX \rangle.$$

10.6. Curvature and the Codazzi–Mainardi equations. We start with the obvious identity

$$\bar{\nabla}_X \bar{\nabla}_Y - \bar{\nabla}_Y \bar{\nabla}_X = \bar{\nabla}_{[X, Y]}$$

between first order differential operators: it is true when applied to functions, hence to vector fields in \mathbb{R}^n .

Substituting into this identity the Gauss identity $\nabla_X = \bar{\nabla}_X - \langle WX, \cdot \rangle N$, applying the result to a third vector field $Z \in \mathcal{X}(\mathbb{R}^n)$ and using several times the Leibniz rule and linearity of the Weingarten operator W , we can separate at the end the normal and tangential components of the result. These will yield us two identities:

$$(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X, Y]})Z = \langle WY, Z \rangle WX - \langle WX, Z \rangle WY, \quad (29)$$

and another identity valid for any Z , which implies that

$$\nabla_X(WY) - \nabla_Y(WX) = W[X, Y]. \quad (30)$$

The identity (29) is very remarkable: it asserts that a certain differential operator that could a priori be of order 2 with respect to Z and of order 1 with respect to X, Y , is of order zero with respect to all arguments! In the classical language, it is a tensor, called the *curvature tensor*, usually considered as a multilinear scalar function of *four*

vector arguments

$$\begin{aligned} R(X, Y, Z, V) &= \langle (\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X, Y]}) Z, V \rangle \\ &= \langle \langle WY, Z \rangle WX - \langle WX, Z \rangle WY, V \rangle \\ &= \langle WY, Z \rangle \langle WX, V \rangle - \langle WX, Z \rangle \langle WY, V \rangle. \end{aligned} \quad (31)$$

Note that the Weingarten operator is self-adjoint, which means that the curvature tensor has a very rich symmetry with respect to permutations of the arguments. The least obvious of these is the so called (the first) *Bianchi formula*, which is obtained by substituting the symmetry assumption $\nabla_X Y - \nabla_Y X = [X, Y]$ into the Jacobi identity $[X, [Y, Z]] + [Y, [X, Z]] + [Z, [X, Y]] = 0$.

Problem 10.1. Write down this identity explicitly.

For the case of 2-surfaces in \mathbb{R}^3 , the geometric nature of this tensor can be described as follows. Let X, Y be two orthonormal vectors. Then

$$\begin{aligned} R(X, Y, X, Y) &= \langle WX, Y \rangle \langle WY, X \rangle - \langle WY, X \rangle \langle WY, Y \rangle \\ &= -\det W = -\lambda_1 \lambda_2 = -R(X, Y, Y, X) \end{aligned} \quad (32)$$

is the Gaussian curvature of the surface.

This formula is important since the left hand side of (29) is defined in terms of the intrinsic geometry of the surface M (the connexion ∇ and the Riemannian metric), while the Weingarten operator depends on the embedding of the surface M into \mathbb{R}^3 .

10.7. How unique is the Riemannian connexion? There is only one covariant derivative which is compatible with a given Riemannian metric on a manifold. Recall that we conveniently denote the Lie derivation on functions as $\nabla_X = L_X = X$.

Theorem 10.2. *The covariant derivation ∇ on a Riemannian manifold, which is symmetric and preserves the Riemannian structure, i.e., $\forall X, Y, Z \in \mathcal{X}(M)$ satisfying the conditions*

$$\nabla_X Y - \nabla_Y X = [X, Y], \quad \nabla_X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle,$$

is unique.

Remark 10.1. The compatibility condition means that the parallel transport along any curve is an isometry between the respective tangent spaces. Indeed, if both $X(t), Y(t)$ are parallel along γ , i.e., $\nabla_{\dot{\gamma}} X =$

$\nabla_{\dot{\gamma}}Y = 0$, then

$$\begin{aligned} \langle X(1), Y(1) \rangle - \langle X(0), Y(0) \rangle &= \int_0^1 \nabla_{\dot{\gamma}} \langle X(t), Y(t) \rangle dt \\ &= \int_0^1 \langle \nabla_{\dot{\gamma}}X, Y \rangle + \langle X, \nabla_{\dot{\gamma}}Y \rangle dt = 0. \end{aligned}$$

Proof of the Theorem. Let $X, Y, Z \in \mathcal{X}(M)$ be any three commuting vector fields. We have the following identities:

$$\begin{aligned} \nabla_X \langle Y, Z \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle, \\ \nabla_Y \langle X, Z \rangle &= \langle \nabla_Y X, Z \rangle + \langle X, \nabla_Y Z \rangle, \\ \nabla_Z \langle X, Y \rangle &= \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle \end{aligned} \tag{33}$$

Adding the first two and subtracting the third identity, we have

$$\begin{aligned} \nabla_X \langle Y, Z \rangle + \nabla_Y \langle X, Z \rangle - \nabla_Z \langle X, Y \rangle &= \\ \langle \nabla_X Y, Z \rangle + \langle Y, [X, Z] \rangle + \langle \nabla_Y X, Z \rangle + \langle X, [Y, Z] \rangle &= \\ = 2 \langle \nabla_X Y, Z \rangle - \langle Z, [X, Y] \rangle = 2 \langle \nabla_X Y, Z \rangle. \end{aligned} \tag{34}$$

Note that the left hand side depends only on the Riemannian metric, while the right hand side involves the covariant derivative $\nabla_X Y$.

Let $E_1, \dots, E_n \in \mathcal{X}(M)$ be coordinate vector fields for any local coordinate system (commuting by definition), and apply the identity (34) to all triples E_i, E_j, E_k . As a result, for any pair i, j one gets an expression of the covariant derivative $\nabla_{E_i} E_j$ via its projections on each direction E_k in terms of the Riemannian metric. This determines the covariant derivative uniquely, if the Riemannian metric is non-degenerate (it always is by definition). \square

This implies what Gauss called *Theorema Egregium* and is the “cartographer’s nightmare”: the surface of the Earth cannot be rendered isometrically on the flat paper. Indeed, the Gauss curvature of the sphere is positive (compute it for the sphere of radius $r > 0$), while that of the plane is zero. Yet the globus is perfect as a scaled image.

References

The classical source is the book [M]. The book [H] was used to explain the Weingarten map and the geometry of submanifolds.

- [M] J. Milnor, *Morse theory*, Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51, Princeton University Press, Princeton, N.J., 1963. MR0163331

- [H] N. J. Hicks, *Notes on differential geometry*, Van Nostrand Mathematical Studies, No. 3, D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1965. MR0179691

11. Geodesics

Let M be a Riemannian manifold with the metric $\langle \cdot, \cdot \rangle = \sum_{i,j} g_{ij}(x) dx_i dx_j$ and the Riemannian connexion ∇ such that $\nabla_{\partial_i} \partial_j = \sum_{k=1}^n \Gamma_{ij}^k(x) \partial_k$, where $\partial_i = \frac{\partial}{\partial x_i}$ in the local coordinates (x_1, \dots, x_n) .

11.1. Some obvious computations. For a vector field $V = (v_1(x), \dots, v_n(x)) = \sum_i v_i(x) \partial_i$ to be parallel along smooth curve $\gamma: t \mapsto x(t)$ with the velocity vector $\dot{\gamma}(t) = (\dot{x}_1(t), \dots, \dot{x}_n(t)) = (w_1(t), \dots, w_n(t))$, $t \in [0, T]$, one has to meet the condition $\nabla_{\dot{\gamma}} V = 0$, i.e.,

$$0 = \nabla_{\dot{\gamma}} \left[\sum_i v_i \partial_i \right] = \sum_i \left[(\nabla_{\dot{\gamma}} v_i) \partial_i + v_i \sum_{k,j} \Gamma_{ij}^k w_j \partial_k \right],$$

This vector equation yields n scalar ODEs

$$\dot{v}_k(t) + \sum_{ij} \Gamma_{ij}^k(t) v_i(t) w_j(t) = 0, \quad k = 1, \dots, n.$$

This can be considered as a system of linear equations for the unknown functions $V(t) = (v_1(t), \dots, v_n(t))$ on $[0, T]$. For any initial condition $V(0)$ it uniquely defines the solution V along any interval. The vector $V(T)$ is the result of the parallel transport along γ from $\gamma(0)$ to $\gamma(T)$.

The requirement that γ is geodesic, takes the form $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$. Then in the above equations $w = v = \dot{x}(t)$ are derivatives of the unknown functions $x(t)$ describing γ . Thus the differential equation for geodesics is

$$\ddot{x}_k(t) = \sum_{i,j} \Gamma_{ij}^k(x(t)) \dot{x}_i(t) \dot{x}_j(t).$$

This is a *second order* nonlinear differential equation whose solution requires the initial conditions

$$x(0) = a \in \mathbb{R}^n, \quad \dot{x}(0) = v \in T_a \mathbb{R}^n \simeq \mathbb{R}^n.$$

Solution of this equation is guaranteed by general theorems of ODE, but in general only locally, for sufficiently small values of t , in general, depending on a and v .

Note that if $\gamma = \gamma_{a,v}(t)$ is a geodesic curve defined on $[0, T]$, then for any $c \in \mathbb{R}$ the curve $\gamma(ct): [0, c^{-1}T]$ is also geodesic with the initial

velocity cv (in particular, $\gamma(t) \equiv a$ is defined on $[0, +\infty)$ and has the initial velocity 0. This allows introduce the self-consistent notation

$$\exp_a(tv): (a, v, t) \longmapsto \gamma_{a,v}(t).$$

It is defined for all combinations (t, v) such that $t\|v\|$ is sufficiently small. This freedom can be used to ensure that the *exponential map* above is defined, say, for all $|t| < 1$.

Proposition 11.1. For every point $a \in M$ on a Riemannian manifold there exists a positive $\varepsilon > 0$ such that the exponential map

$$\exp: TM \times \mathbb{R} \rightarrow M, \quad (a, v, t) \longmapsto \gamma_{a,v}(t),$$

is well defined in the ε -neighborhood of the point $(a, 0) \in TM$ for all $|t| < 1$.

In other words, from any point $a \in M$ and in any direction $v \in T_aM$ one can find a geodesic curve $\gamma_{a,v}: t \mapsto M$ passing through this point and tangent to this direction.

Computing the first order terms of \exp in t , we see that

$$\exp_a(tv) = a + tv + o(|t|).$$

Combining this with the inverse function theorem, we prove the following theorem.

Theorem 11.1. *The exponential map*

$$v \longmapsto \exp_a(v)$$

is a diffeomorphism between a small enough neighborhood of the origin in T_aM and a small neighborhood of a in M , which sends lines $\{tv : t \in \mathbb{R}\}$ into geodesic curves.

For compact Riemannian manifolds (without boundaries) these local results suffice to prove what Euclid had to postulate for straight lines: any geodesic curve on any compact manifold can be extended unlimited.

For noncompact manifolds one has to explicitly require the *geodesic completeness* which, fortunately, coincides with the topological completeness.

Theorem 11.2 (Hopf–Rinow, 1931). *If a Riemannian manifold is a complete metric space with respect to the metric induced by the Riemannian structure, then it is geodesically complete, i.e., \exp_a is well defined on the whole T_aM for any $a \in M$, and vice versa.*

11.2. Local minimality. Consider a point $a \in M$ and the exponential map $\exp_a: T_aM \rightarrow (M, a)$. Let $S: (M, a) \rightarrow \mathbb{R}_+$ be the transplant of the squared norm: $S(\exp_a(v)) = \langle v, v \rangle$. This is a smooth function on M near a , whose level hypersurfaces are images of the spheres $\langle v, v \rangle = \text{const}$. The rays $\mathbb{R}_+v = \{rv: v \in T_aM, r \in \mathbb{R}_+\} \subset T_aM$ are orthogonal to the spheres.

The exponential map is not an isometry, however, the above orthogonality is preserved by the exponentiation.

Proposition 11.2. The geodesics $\gamma_{a,rv} = \exp_a(rv)$ are orthogonal to the level hypersurfaces $S = \text{const}$.

Proof. Consider the vector field R of unit length, tangent to the geodesic pencil, and let X be an arbitrary vector field such that $L_X S = 0$, which commutes with R , $[R, X] = 0$. We prove that $\langle R, X \rangle \equiv 0$. For that sake differentiate the identity $\langle R, R \rangle \equiv 1$ along X :

$$0 = \nabla_X \langle R, R \rangle = \langle R, \nabla_X R \rangle = \langle R, \nabla_R X \rangle,$$

since ∇ is symmetric and $[X, R] = 0$. This implies that “radial derivative” of the scalar product $\langle R, X \rangle$ also vanishes:

$$\nabla_R \langle R, X \rangle = \langle \nabla_R R, X \rangle + \langle R, \nabla_R X \rangle = 0 + 0 = 0,$$

since the field R is the velocity of geodesics. Thus $\langle R, X \rangle$ is constant along each geodesic, and obviously in the limit $r \rightarrow 0^+$ we have the orthogonality, so that this constant is zero. Since X was arbitrary, R is orthogonal to $\{S = \text{const}\}$ at every point of the latter. \square

Corollary 11.1. *For as long as the exponential map $\exp_a: T_aM \rightarrow (M, a)$ remains a diffeomorphism, the geodesic curve is “the shortest” smooth curve among all curves with the same points (more precisely, its length is less or equal than the length of any other curve).*

Proof. Let $r(x) = \sqrt{S(x)}$ be the “geodesic distance function” on M from the given point a . The geodesic curves parameterized by the arclength are orthogonal to the surfaces $r = \text{const}$. Therefore for any smooth curve the absolute value of the differential dr on the velocity vector is no greater than the norm of the velocity vector (in the same way as the leg of an orthogonal triangle cannot be greater than its hypotenuse). Integrating the norm of velocity, we see that the length of any curve connecting a and b is no greater than $r(b)$, the length of the geodesic curve from a to b . \square

12. Lie groups and Lie algebras

12.1. Definitions. A smooth manifold G is called a Lie group, if it carries the group structure (i.e., points can be multiplied between themselves according to a group law), and for each element $g \in G$ the application $G \times G \rightarrow G$, $(g, h) \mapsto gh^{-1}$ is smooth. One can instantly see that in this case the maps $l_g: h \mapsto gh$, $r_g: h \mapsto hg$ and $h \mapsto h^{-1}$ are smooth self-maps of G for any g . They are called *left shift*, *right shift* and *reciprocal involution* of G .

The unit of the group will be denoted by e and it obviously will be a very distinguished point of G .

Lie groups can be commutative or not, and usually we will assume them to be connected. For non-connected group, the component G_e containing e is a normal subgroup and the quotient G/G_e is a discrete (zero-dimensional) group (Prove that!).

12.2. Examples. Discrete groups (zero-dimensional) are trivial and not interesting for us.

We list the principal examples of (finite-dimensional) Lie groups⁴.

- (1) Linear spaces \mathbb{R}^n with the usual structure of Abelian additive group.
- (2) The multiplicative groups $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ and the unit circle $T^1 = \{|z| = 1\} \subseteq \mathbb{C}^*$, also Abelian.
- (3) Tori $\mathbb{R}^n/\mathbb{Z}^n$, also Abelian.
- (4) The *general linear group* $GL(n, \mathbb{R})$ of real nondegenerate $n \times n$ -matrices.
- (5) The *special linear group* $SL(n, \mathbb{R}) = \{M \in GL(n, \mathbb{R}) : \det M = 1\}$ of matrices with the unit determinant.
- (6) The *orthogonal group* $O(n, \mathbb{R}) = \{M \in GL(n, \mathbb{R}) : \langle Mx, My \rangle = \langle x, y \rangle\}$ of orthogonal matrices preserving the Euclidean structure on \mathbb{R}^n (rigid rotations around the origin and reflections).
- (7) The *special orthogonal group* $SO(n, \mathbb{R}) = \{M \in O(n, \mathbb{R}) : \det M = 1\}$ orientation-preserving rotations.

Basically, the group of linear operators preserving any algebraic structure on \mathbb{R}^n (symplectic, complex for even n etc.) is a Lie group. E.g., upper-triangular nondegenerate matrices preserve a complete flag of subspaces in \mathbb{R}^n .

12.3. Constructions. The standard group-theoretic constructions yield more Lie groups.

⁴Infinite-dimensional Lie groups are also immensely important.

- (1) Cartesian (direct) product.
- (2) Semidirect product (amalgam). E.g., the group $\text{SO}(n, \mathbb{R}) \ltimes \mathbb{R}^n$ of all rigid orientation-preserving transformations of a Euclidean space. The explicit formula for the composition $(M, v) \cdot (N, w)$ can be derived from the formula $(M, v) \cdot x = Mx + v$ for all $x \in \mathbb{R}^n$.
- (3) Quotient groups G/G' in the case where G' is a normal discrete subgroup in G . Example: $\text{SO}(n, \mathbb{R}) = \text{O}(n, \mathbb{R})/\{\pm E\}$.

12.4. Invariant vector fields and Lie algebras. Unlike general manifolds, the Lie groups naturally come equipped with a flat connection (actually, two of them). By definition, any Lie group G acts on itself by the left shifts $l_g: h \mapsto gh$. Such shift takes $e \in G$ into $g \in G$, moreover, any two points $g, h \in G$ can be transformed one into the other by the unique left shift, $l_{hg^{-1}}(g) = h$. The differential $dl_{hg^{-1}}: T_g G \rightarrow T_h G$ provides the parallel transport between the tangent spaces at these points, which is independent of any curve connecting g with h .

Thus for any vector $v \in T_e G$ tangent to G at the unit point of the group can be carried out to any point g by the differential $dl_g(v) \in T_g G$, together forming the vector field X which is *left invariant*, $dl_g X = X$ for all $g \in G$. Note that there are no nontrivial left-invariant functions on G except for constants (why?).

One can instantly see that the property of being left invariant is preserved by the algebraic operations (linear combinations) and commutator. The flow of any invariant vector field commutes with any left shift.

Conversely, any left invariant vector field X is uniquely defined by its value $X(e) \in T_e G$. Thus the space of all invariant vector fields is finite-dimensional and isomorphic to $T_e G$. The operator of commutator becomes then a bilinear non-commutative operator on $T_e G$, satisfying the standard conditions

$$[X, Y] = -[Y, X], \quad [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]]. \quad (35)$$

Definition 12.1. A finite-dimensional linear space \mathfrak{g} equipped with a bilinear operation (35), is called the Lie algebra. The operation is called the *Lie bracket*.

The above construction with invariant vector fields shows that each Lie group G is uniquely associated with a Lie algebra.

12.5. Examples. Any vector space can be turned into a Lie algebra with the trivial (null) Lie bracket. Such Lie algebras, called *commutative*, appear as Lie algebras of Abelian Lie groups.

The group $\mathrm{GL}(n, \mathbb{R})$ is an open subset of the linear space $\mathbb{R}^{n^2} = \mathrm{Mat}(n, \mathbb{R})$, so its Lie algebra $\mathfrak{gl}(n, \mathbb{R})$ consists of all matrices. The Lie bracket is given by the matrix commutator $[X, Y] = XY - YX$.

For all other subgroups the Lie algebras will be linear subspaces of $\mathfrak{gl}(n, \mathbb{R})$ with the commutator as the Lie bracket. To compute then it suffices to find these subspaces explicitly.

Since $\det(E + tX) = 1 + t \cdot \mathrm{tr} X + o(t)$, the special linear Lie algebra $\mathfrak{sl}(n, \mathbb{R}) = \{X : \mathrm{tr} X = 0\}$ consists of traceless matrices.

In turn, if $E + tX + \dots$ is orthogonal, $(E + tX + \dots)(E + tX^* + \dots) = E$, the special orthogonal Lie algebra $\mathfrak{so}(n, \mathbb{R}) = \{X : X^* = -X\}$ consists of antisymmetric matrixes.

Let $X_1, \dots, X_n \in \mathfrak{g}$, $n = \dim G = \dim \mathfrak{g}$ be a basis in the Lie algebra. Then the Lie bracket is completely determined by the collection of real numbers $\{c_{ijk}\} \in \mathbb{R}^{3n}$ such that

$$[X_i, X_j] = \sum_{k=1}^n c_{ijk} X_k.$$

These numbers are called the *structural constants*, and they satisfy the condition $c_{ijk} + c_{jik} = 0$ and another one, implied by the Jacobi identity.

12.6. Curvature and torsion of the Lie connection. The curvature of the connection induced on the Lie group, is zero. If $X, Y \in \mathfrak{g}$ are two invariant vector fields, then their integral trajectories are geodesic, moreover, $\nabla_X Y = 0$. Thus the torsion $\nabla_X Y - \nabla_Y X - [X, Y]$ coincides with the Lie bracket up to the sign.

12.7. Maurer–Cartan forms. In the same way using the left shifts, one can identify all cotangent spaces T^*gG with the cotangent space at the unity of the group and introduce the n -dimensional space of left invariant 1-forms from $\Lambda^1(G)$, called the *Maurer–Cartan forms*, dual to \mathfrak{g} . This construction immediately generalizes for invariant k -forms.

For an invariant form ω and invariant field X the value $\omega(X)$ is an invariant function, i.e., a constant. Thus

$$d\omega(X, Y) = -\omega([X, Y]).$$

Choosing a basis of invariant 1-forms $\omega_1, \dots, \omega_n$, one can express their exterior derivatives as linear combinations,

$$d\omega_i = \sum_{j < k} c'_{ijk} \omega_j \wedge \omega_k.$$

One can easily express the structural constants c'_{ijk} via c_{ijk} in the case where ω_i are dual to X_j as above, $\omega_i(X_j) = \delta_{ij}$.

12.8. Interplay between G and \mathfrak{g} . Morally, \mathfrak{g} carries only linearization data about the group structure on G . However, it turns out that these data are sufficient to reconstruct precisely large enough piece of G from \mathfrak{g} . In particular, for connected simply connected Lie groups \mathfrak{g} uniquely defines G . However, globally this no longer holds: compare various Abelian Lie groups with isomorphic (commutative) Lie algebras, in particular, tori with the Euclidean spaces.

Example 12.1. Any Lie subgroup $H \subseteq G$ is associated with a Lie subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$. The inverse in general is not true, or, more precisely, depends on the topology. Consider the rational and irrational windings of the torus.

Theorem 12.1. *Let G, H be two Lie groups with the Lie algebras $\mathfrak{g}, \mathfrak{h}$ respectively, and assume that G is simply connected. Then any Lie algebra homomorphism $\phi: \mathfrak{g} \rightarrow \mathfrak{h}$ extends to a Lie group homomorphism (smooth map preserving the group structure) $f: G \rightarrow H$ such that $df = \phi$.*

12.9. The exponential map. Let $X \in \mathfrak{g}$ be an element of the Lie algebra of G , interpreted as a left invariant vector field. Then the trajectory of this vector field through the unit is a map $t \mapsto g(t)$, such that $g(0) = e$, $\dot{g}(0)e = X(e)$. One can instantly see that the invariance of X implies that $g(t+s) = g(t)g(s)$. That is, $g(\cdot)$ is a homomorphism of the commutative Lie group \mathbb{R} to G , which sends $\frac{d}{dt}$ into $X(e)$, so that we have $\dot{g}(t) = dl_{g(t)}X(e)$.

Definition 12.2. The exponential map $\exp: \mathfrak{g} \rightarrow G$ is the map that sends each vector $X \in T_eG \simeq \mathfrak{g}$ into the point $\exp X = g(1)$, where $g(\cdot): \mathbb{R} \rightarrow G$ is the above one-parameter subgroup.

Theorem 12.2. *The exponential map is a local diffeomorphism between $(\mathfrak{g}, 0)$ and (G, e) , tangent to the identity.*

However, globally the exponential map may cease to be a diffeomorphism. Example: $\exp: \mathbb{R}^1 \rightarrow \mathbb{T}^1$, $\exp x = e^{2\pi i x}$.

For the matrix group $GL(n, \mathbb{R})$ one can explicitly find a one-parametric subgroup tangent to the given vector $X \in \mathfrak{gl}(n, \mathbb{C})$ at the “origin” (the unit matrix E),

$$\exp tX = \sum_{k=0}^{\infty} \frac{t^k}{k!} X^k.$$

It converges for all $t \in \mathbb{R}$ and all $X \in \mathfrak{gl}(n, \mathbb{R})$ and $\exp(t+s)X = \exp tX \cdot \exp sX$ since the matrices sX and tX commute with each other and can be substituted into the formal identity $e^{tx}e^{sx} = e^{(t+s)x}$. The

last identity implies that $\frac{d}{dt}\big|_{t=0} \exp tX = X$, that is, the exponential series indeed is tangent to the prescribed velocity vector X .

References

- [W] F. W. Warner, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, vol. 94, Springer-Verlag, New York-Berlin, 1983. Corrected reprint of the 1971 edition. MR722297

13. Symplectic manifolds and Hamiltonian dynamics

13.1. Symplectic structure. Let N be a smooth manifold and T^*N is cotangent bundle, the union of all duals to tangent spaces $T^*N = \bigcup_{x \in N} T_x^*N$. This bundle has a natural projection $\pi: T^*N \rightarrow N$ on the base. In the local coordinates x_1, \dots, x_n the coordinates on T_x^*N are denoted by p_1, \dots, p_n so that $p_i(\partial/\partial x_j) = \delta_{ij}$. By construction, $\dim T^*N = 2 \dim N$ is always an even number $2n$.

For physical reasons we will refer to N as the *configuration space* and M as the *phase space*. One should imagine the coordinates x_i as the spatial coordinates of a point (or a more complicated mechanical systems) and p_i as the respective moments.

The phase space $M = T^*N$ is naturally equipped with a special 1-form α . Indeed, let $\pi: M \rightarrow N$ be the above projection. Then for any point $a = (x, \xi) \in M$, $\xi \in T_xN$, the differential $d\pi: T_aM \rightarrow T_aN$ is well defined and takes any vector tangent to M at a into a vector tangent to N at x . Then one can evaluate the ξ -component of the “point” a of the phase space on this image vector tangent to the configuration spaces. This defines a 1-form.

In the coordinates the above verbose definition is described by a much simpler expression:

$$\alpha = \sum_{i=1}^n p_i dx_i.$$

Definition 13.1. The *canonical symplectic structure* on T^*M is the differential 2-form $\omega = d\alpha$. In the local canonical coordinates,

$$\omega = \sum_{i=1}^n dp_i \wedge dx_i. \quad (36)$$

One can immediately verify that the 2-form ω on M is *nondegenerate* in the following sense:

$$\forall v \in T_a M \exists w \in T_a M \text{ such that } \omega(v, w) \neq 0.$$

This is the same as to say that the (antisymmetric) matrix of ω in any basis in $T_a M$ is nondegenerate, i.e., has nonvanishing determinant.

Definition 13.2 (main). A symplectic manifold is a manifold equipped with a nondegenerate 2-form ω , called symplectic structure.

Remark 13.1. Since $\det(-A) = (-1)^n \det A$, an odd-sized antisymmetric matrix cannot be nondegenerate, hence a symplectic manifold must be even-dimensional.

Example 13.1. The simplest example of a symplectic manifold is the plane with the (oriented) area as the symplectic structure.

13.2. Hamiltonian vector fields. In the same way as the symmetric Riemannian structure allows to identify vectors with covectors and make a differential of function df , born as 1-form, into the vector field $\text{grad } f$ (note, however, that not any vector field is a gradient), the symplectic structure allows to do the same.

Definition 13.3. A vector field $X \in \mathcal{X}(M)$ on a symplectic manifold is called *locally Hamiltonian*, if $i_X \omega = \omega(X, \cdot) \in \Omega^1(M)$ is closed. The field is called *Hamiltonian*, if $i_X \omega = dH$ is exact. The function $H \in \mathcal{C}(M)$ is called the *Hamiltonian* of the field X .

Thus Hamiltonian vector fields are in a sense “skew-gradients” of the respective Hamiltonians. From the computational point of view, they are much easier to deal with: a Hamiltonian vector field is defined by one function of $2n$ arguments, rather than by $2n$ such functions. But, of course, the primary importance of the Hamiltonian vector fields roots in the fact that the differential equations of the (frictionless) celestial dynamics are Hamiltonian.

13.3. Conservation laws.

Theorem 13.1. *The flow of any locally Hamiltonian vector field preserves the symplectic structure: if $di_X \omega = 0$, then $L_X \omega = 0$.*

Conversely, if $L_X \omega = 0$, then X is locally Hamiltonian.

Proof.

$$L_X \omega = i_X d\omega + di_X \omega = 0. \quad \square$$

The n th power $\omega \wedge \cdots \wedge \omega$ is the (oriented) volume form on M .

Corollary 13.1. *The flow of any locally Hamiltonian vector field is volume-preserving. \square*

Theorem 13.2. *The Hamiltonian function H is preserved by the corresponding Hamiltonian flow.*

Proof.

$$L_X H = i_X dH = i_X i_X \omega = \omega(X, X) = 0. \quad \square$$

13.4. **Poisson brackets.** Obviously, the commutator of two locally Hamiltonian vector fields is locally Hamiltonian:

$$L_{[X,Y]}\omega = (L_X L_Y - L_Y L_X)\omega = 0 - 0 = 0.$$

A simple computation yields for two Hamiltonian vector fields $X = \text{sgrad } H$, $Y = \text{sgrad } G$,

$$\omega(X, Y) = i_X i_Y \omega = L_X G = -L_Y H.$$

Definition 13.4. The above function is called the *Poisson bracket* of H , G and denoted by $\{H, G\}$.

The elementary computation yields

$$\text{sgrad}\{H, G\} = [\text{sgrad } H, \text{sgrad } G].$$

This implies that the Poisson bracket satisfies the Jacobi identity,

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0.$$

Example 13.2. The canonical coordinates (x_i, p_i) considered as functions on the standard symplectic manifold, have the following Poisson brackets:

$$\{x_i, x_j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{x_i, p_j\} = \delta_{ij}.$$

Theorem 13.3. *The algebra of functions $\mathcal{C}(M)$ on a symplectic manifold, is a Lie algebra with respect to the Poisson bracket. The corresponding Hamiltonian vector fields form a Lie subalgebra of the Lie algebra of vector fields $\mathcal{X}(M)$ with the commutator $[\cdot, \cdot]$ as the Lie bracket. The map $H \mapsto \text{sgrad } H$ is a homomorphism of Lie algebras, whose kernel consists of constants, if M is connected.*

13.5. **First integrals.** A function f is called the *first integral* of a vector field X , if $Xf = L_X f = 0$. If a vector field on an m -dimensional manifold has $m - 1$ first integrals f_1, \dots, f_{m-1} with linear independent differentials near a point a , then the field is *integrable*: its trajectories are defined by the common level curves $\{f_1 = c_1, \dots, f_{m-1} = c_{m-1}\}$. Locally the converse is true.

A Hamiltonian vector field $X = \text{sgrad } H$ on a symplectic manifold (M, ω) always has at least one first integral H . The Jacobi identity

means that the Poisson bracket of two first integrals is a first integral again.

13.6. Symplectic geometry. A real linear even-dimensional space \mathbb{R}^{2n} equipped with a non-degenerate antisymmetric bilinear form $[\cdot, \cdot]$, is called the symplectic space. We abuse the term and say that two vectors are *orthogonal*, $v \perp w$, if $[v, w] = 0$. Note that, unlike in the symmetric case, every vector is orthogonal to itself. The non-degeneracy condition means that there is no vector orthogonal to the entire symplectic space, so the orthogonal complement v^\perp is always $(2n - 1)$ -dimensional.

A linear map between the symplectic spaces into itself is called *symplectic*, if it respects the corresponding structure. In particular, a self-map S is symplectic if and only if $[Sv, Sw] = [v, w]$ for any two vectors.

Theorem 13.4. *Any symplectic space of even dimension $2n$ always has a basis*

$$v_1, \dots, v_n, w_1, \dots, w_n$$

such that

$$[v_i, v_j] = [w_i, w_j] = 0, \quad [v_i, w_j] = \delta_{ij}.$$

Therefore all symplectic spaces of the same dimension are isomorphic.

References

- [A] V. I. Arnold, *Mathematical methods of classical mechanics*, 2nd ed., Graduate Texts in Mathematics, vol. 60, Springer-Verlag, New York, 1989. Translated from the Russian by K. Vogtmann and A. Weinstein. MR997295

14. Complex manifolds

14.1. A philosophical remark on complex numbers. The field \mathbb{C} of complex numbers is obtained by an algebraic extension: the root(s) of the equation $x^2 + 1 = 0$ are added to the field of reals \mathbb{R} . It turns out that \mathbb{C} is both topologically complete and algebraically closed. There were absolutely no reasons for those two very useful properties coincide. Although the field \mathbb{C} cannot be ordered, it contains the ordered subfield \mathbb{R} which adds to the richness of the corresponding construction.

We denote the newly added numbers by $\pm i$, but it impossible to distinguish between these Siamese twins: the swap $i \mapsto -i$ induces the (Galois) automorphism of \mathbb{C}/\mathbb{R} , called the complex conjugacy:

$$\overline{x + iy} = x - iy, \quad x, y \in \mathbb{R}.$$

The conjugacy is continuous, keeps all field operations in \mathbb{C} and is identical on \mathbb{R} .

Any polynomial from $\mathbb{C}[x, y]$ can be rewritten as a polynomial from two *independent* variables $z = x + iy$ and $\bar{z} = x - iy$, using the transformation $x = \frac{1}{2}(z + \bar{z})$, $y = \frac{i}{2}(z - \bar{z})$. If the polynomial is real, $p \in \mathbb{R}[x, y]$, then its rendering in $\mathbb{C}[x, y]$ is *symmetric*: $\overline{p(\bar{z}, z)} = p(z, \bar{z})$. Indeed, if $c_{ij} = \bar{c}_{ij}$, then

$$\sum c_{ij} (z + \bar{z})^i \overline{(i(z - \bar{z}))^j} = \sum c_{ij} (\bar{z} + z)^i (i(\bar{z} - z))^j = p(\bar{z}, z).$$

Polynomials which depend only on z and not on \bar{z} , are holomorphic functions on \mathbb{C} .

14.2. Linear theory. The spaces \mathbb{C}^n and \mathbb{R}^{2n} are isomorphic as linear \mathbb{R} -spaces. From the “real” point of view the complex structure (multiplication by i) is an \mathbb{R} -linear operator $J: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ such that $J^2 = -\text{id}$. Presence of such operator allows to define the complex multiplication

$$\mathbb{C} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}, \quad (a + ib)v = av + bJv, \quad v \in \mathbb{R}^{2n}.$$

A map $A: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ is \mathbb{C} -linear, if and only if it commutes with J , $AJ = JA$. A tuple of vectors $(v_1, \dots, v_n) \in \mathbb{R}^{2n}$ is a \mathbb{C} -basis of this space, if $\{v_1, \dots, v_n, Jv_1, \dots, Jv_n\}$ is the \mathbb{R} -basis of \mathbb{R}^{2n} .

This process can be applied to a \mathbb{C} -space $V \simeq \mathbb{R}^{2n}$. Considered as an \mathbb{R} -linear space $V^{\mathbb{R}}$, it has a \mathbb{R} -linear self-map J with the above property, which in a suitable \mathbb{R} -basis has the matrix

$$J = \begin{pmatrix} & E \\ -E & \end{pmatrix}, \quad E \text{ the identity matrix.}$$

The operator J has $2n$ eigenvalues $\pm i$, each with multiplicity n , but is not diagonal, neither can it be diagonalized without additional *complexification*: one has pass from $V^{\mathbb{R}}$ to the linear space $V^{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C} = V^{\mathbb{R}\mathbb{C}}$. The latter is a \mathbb{C} -space of complex dimension $\dim_{\mathbb{C}} V^{\mathbb{R}\mathbb{C}} = 2n = \dim_{\mathbb{R}} V^{\mathbb{R}} = 2 \dim_{\mathbb{C}} V$.

Example 14.1. Let $n = 1$. Then $V = \mathbb{C} \simeq \mathbb{R}^2$ and in the coordinates (x, y) such that $z = z + iy$ we have $J = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$. To diagonalize J , we allow the coordinates to take the *complex* rather than real values,

and introduce two new coordinates $x = \frac{1}{2}(z + w)$, $y = \frac{i}{2}(z - w)$, so that $z = x + iy$ and $w = x - iy$. In the new coordinates the matrix of J becomes diagonal, $J = \begin{pmatrix} i & \\ & -i \end{pmatrix}$.

Remark 14.1. This duplication seemingly contradicting the Occam razor principle, can be explained rather easily: once we forgot for a moment the original complex multiplication on V , we cannot distinguish multiplication by i from the multiplication by its conjugate $\bar{i} = -i$, thus any unification of real and complex structures must be invariant with respect to the automorphism $\alpha \mapsto \bar{\alpha}$ of \mathbb{C} .

Remark 14.2. When passing from the real coordinates (x, y) on $V^{\mathbb{R}}$ to the complex coordinates on $V^{\mathbb{R}\mathbb{C}}$, it is natural to call the first coordinate again z . The other coordinate w is conveniently denoted by \bar{z} .

Thus if we start with the complex space $V = \mathbb{C}^n$ with the coordinates z_1, \dots, z_n , $z_i \in \mathbb{C}$, the result $V^{\mathbb{R}\mathbb{C}} = \mathbb{C}^{2n}$ is naturally equipped with the coordinates $z_1, \dots, z_n, \bar{z}_1, \dots, \bar{z}_n$. These are *independent* complex coordinates which take the conjugate values only on the *real* subspace $\mathbb{R}^n \subset \mathbb{C}^n$.

The structure induced by the operator J on the even-dimensional space $V^{\mathbb{R}\mathbb{C}}$, makes subspaces L of this space unequal. Recall that the eigenvalues of J are $\pm i$. This splits $V^{\mathbb{R}\mathbb{C}}$ into the direct sum of two eigenspaces, $V^{\mathbb{R}\mathbb{C}} = V' + V''$, where

$$V' = \{v \in V^{\mathbb{R}\mathbb{C}} : Jv = iv\}, \quad V'' = \{v \in V^{\mathbb{R}\mathbb{C}} : Jv = -iv\}.$$

Obviously, $\dim_{\mathbb{C}} V', V'' = n$. The first subspace can naturally be identified with the initial \mathbb{C} -space V , the other is called the conjugate.

There above splitting induces another *involution* operation on $V^{\mathbb{R}\mathbb{C}}$, also called complex conjugacy (extending the conjugacy $(a + ib) = z \mapsto \bar{z} = a - ib$) and denoted by the bar, $v \mapsto \bar{v}$. This operator is \mathbb{R} -linear, but \mathbb{C} -*antilinear*: $\overline{\bar{v}} = \bar{\bar{v}} = -i\bar{v}$. In the coordinates introduced above, the conjugacy exchanges

$$\bar{\cdot} : V^{\mathbb{R}\mathbb{C}} \rightarrow V^{\mathbb{R}\mathbb{C}}, \quad (z, \bar{z}) \mapsto (\bar{z}, z).$$

This notation allows to write the above splitting as $V^{\mathbb{R}\mathbb{C}} = V + \bar{V}$.

14.3. Hermitian forms. There is no suitable definition of positivity over \mathbb{C} (no way to introduce a complete order compatible with the natural axioms), so it is impossible to introduce a properly behaving scalar product (the Euclidean structure).

The nearest best choice is the Hermitian structure, which involves the complex conjugacy. It capitalizes on the fact that $z\bar{z} \geq 0$ (and positive

for $\mathbb{C} \ni z \neq 0$). This is the pairing $\mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$, which is *sesquilinear* bilinear over \mathbb{R} , \mathbb{C} -linear in the first argument but antilinear in the second argument,

$$\langle \alpha u, v \rangle = \alpha \langle u, v \rangle, \quad \langle u, \alpha v \rangle = \bar{\alpha} \langle u, v \rangle.$$

The model example is given by the formula $\langle u, v \rangle = \sum_{i=1}^n u_i \bar{v}_i$. Any other *sesquilinear* form that is positive, $\langle u, u \rangle \geq 0$, $\langle u, u \rangle = 0 \iff u = 0 \in \mathbb{C}^n$, is equivalent to it.

It is an important fact that the real part (\cdot, \cdot) of the Hermitian structure is \mathbb{R} -bilinear and symmetric, hence a Euclidean structure (genuine scalar product, positive on nonzero vectors), while the imaginary part $[\cdot, \cdot]$ is antisymmetric and nondegenerate, hence a symplectic structure on \mathbb{C}^n ,

$$\operatorname{Re} \langle u, v \rangle = (u, v), \quad \operatorname{Im} \langle u, v \rangle = [u, v].$$

14.4. Dual spaces and their decomposition. For an \mathbb{R} -linear even-dimensional space \mathbb{R}^{2n} the usual dual, the space of \mathbb{R} -linear maps $\mathbb{R}^{2n} \rightarrow \mathbb{R}$, is isomorphic to \mathbb{R}^{2n*} . If V is a \mathbb{C} -linear space, forgetting the complex structure we obtain the $\operatorname{Hom}_{\mathbb{R}}(V, \mathbb{R})$. By tensoring out with \mathbb{C} we construct the complexification of the latter space.

Abusing notation, we will use the star notation to denote for a \mathbb{C} -linear vector space V this complexified dual,

$$V^* = \operatorname{Hom}_{\mathbb{R}}(V, \mathbb{C}) = \operatorname{Hom}_{\mathbb{R}}(V, \mathbb{R}) \otimes_{\mathbb{R}} \mathbb{C}.$$

This larger space naturally splits into \mathbb{C} -linear and \mathbb{C} -antilinear maps $\xi: V \rightarrow \mathbb{C}$,

$$V_{1,0}^* = \{\xi(\alpha v) = \alpha \xi(v)\}, \quad \text{resp.}, \quad V_{0,1}^* = \{\xi(\alpha v) = \bar{\alpha} \xi(v)\},$$

(suffices to require these conditions for $\alpha = i$). Note that V^* is a \mathbb{C} -linear space, and $\dim_{\mathbb{C}} V^* = 2n = 2 \dim_{\mathbb{C}} V = \dim_{\mathbb{R}} V$.

If we equip \mathbb{C}^n with the coordinates z_1, \dots, z_n , $x_i + iy_i = z_i \in \mathbb{C}$, then the (constant) forms $dz_i = dx_i + idy_i$ and $d\bar{z}_i = dx_i - iy_i$ constitute the basis of \mathbb{C}^{n*} .

This splitting is naturally dual to the splitting of $V^{\mathbb{R}\mathbb{C}} = V + \bar{V}$.

14.5. Higher exterior powers. We can form the higher exterior powers $\bigwedge^k V^*$, $k = 1, \dots, 2n$, as multi- \mathbb{R} -linear antisymmetric forms on V with values in \mathbb{C} . The decomposition $V^* = V_{1,0}^* \oplus V_{0,1}^*$ yields then the decomposition

$$\bigwedge^k V^* = \sum_{p+q=k} V_{p,q}^*,$$

where $V_{p,q}^*$ is the space spanned over \mathbb{C} by *monomials* of the form $\xi_1 \wedge \cdots \wedge \xi_p \wedge \eta_1 \wedge \cdots \wedge \eta_q$ with $\xi_i \in V_{1,0}^*$, $\eta_j \in V_{0,1}^*$.

14.6. Holomorphic functions and maps. A map $f: U \rightarrow \mathbb{C}^m$, $U \subseteq \mathbb{C}^n$, is holomorphic at a point $a \in U$, if its differential (tangent map) $df: \mathbb{C}^n \rightarrow \mathbb{C}^m$ is \mathbb{C} -linear. The map is holomorphic in U , if it is holomorphic at every point of U . The map is a biholomorphism, if its inverse is also holomorphic at all points of the image $f(U)$. The set of all \mathbb{C} -linear self-maps of \mathbb{C}^n into itself is a group (denoted by $GL(n, \mathbb{C})$). Composition of holomorphic maps is again holomorphic. The space of holomorphic maps $U \rightarrow \mathbb{C}$ (holomorphic functions) is denoted by $\mathcal{O}(U)$. A basic theorem of complex analysis claims that $\mathcal{O}(U) \subset \mathcal{C}(U)$, that is, existence of “just one derivative” (\mathbb{C} -linearity of df) implies infinite differentiability of the latter.

14.7. Holomorphic manifolds and complex differential forms on them. A *holomorphic manifold* is a topological space M equipped with an atlas of charts (homeomorphisms) $z_\alpha: U_\alpha \rightarrow \mathbb{C}^n$ so that $M = \bigcup_\alpha U_\alpha$ and all transition maps $z_\alpha \circ z_\beta^{-1}$ are biholomorphisms.

The tangent spaces $T_a M$ to M at different points $a \in M$ are naturally isomorphic to \mathbb{C}^n . The *complex cotangent space* T_a^* is defined as the above complex dual space, $T_a^* M = \text{Hom}_{\mathbb{R}}(T_a M, \mathbb{C})$, $\dim_{\mathbb{R}} T_a^* M = 2 \dim_{\mathbb{C}} T_a^* M = 4n$.

A (smooth complex) differential 1-form is a C^∞ -smooth map which selects $\xi(a) \in T_a^* M$ for each a . By this definition, locally any such form can be represented as

$$\xi = \sum_{i=1}^n a_i(z) dz_i + b_i(z) d\bar{z}_i, \quad a_i, b_i \in \mathcal{C}(U)$$

with smooth complex-valued coefficients. To stress the fact that these coefficients are not necessarily holomorphic, sometimes we write $a_i(z, \bar{z})$, $b_i(z, \bar{z})$.

Respectively, we have the $\mathcal{C}(U)$ -module $\Omega^{p,q}(U)$ of forms of the type (p, q) , locally representable as

$$\sum_{i_1 < \cdots < i_p, j_1 < \cdots < j_q} a_{i,j}(z) dz_{i_1} \wedge \cdots \wedge dz_{i_p} \wedge d\bar{z}_{j_1} \wedge \cdots \wedge d\bar{z}_{j_q}.$$

One can verify that if a form is of type (p, q) in some chart, then it is of the same type in any other chart. This follows from the fact that the splitting into holomorphic and antiholomorphic forms is preserved by holomorphic maps.

14.8. How do we differentiate? The operator of exterior derivation is uniquely defined by its values on functions, and we don't have much freedom if we want it to vanish on the complex constants as well as on the real ones. Thus if $f(z) = u(z) + iv(z) \in \mathcal{C}(U)$, then $df = du + idv$, if $u, v: U \rightarrow \mathbb{R}$ are C^∞ -smooth real and imaginary parts of a complex valued function $f: U \rightarrow \mathbb{C}$.

However, the splitting into the real and imaginary parts is not as convenient, as the splitting into the holomorphic and antiholomorphic parts.

Definition 14.1. Consider the following differential operators on $\mathcal{C}(U)$, which in the coordinates (z_1, \dots, z_n) , $z_i = x_i + iy_i$, take the form

$$\partial_i = \frac{\partial}{\partial z_i} = \frac{1}{2} \left(\frac{\partial}{\partial x_i} - i \frac{\partial}{\partial y_i} \right), \quad \bar{\partial}_i = \frac{\partial}{\partial \bar{z}_i} = \frac{1}{2} \left(\frac{\partial}{\partial x_i} + i \frac{\partial}{\partial y_i} \right).$$

The operators $\partial_i, \bar{\partial}_i$ are Leibniz-respecting \mathbb{C} -linear operators $\mathcal{C}(U)$ to itself, so can be considered as vector fields on U .

Proposition 14.1 (Cauchy–Riemann conditions). A function $f \in \mathcal{C}(U)$ is holomorphic, $f \in \mathcal{O}(U)$, if and only if $\bar{\partial}_i f = 0$ for all $i = 1, \dots, n$.

Proposition 14.2. The forms dz_i and $d\bar{z}_i$ are dual to the vector fields $\partial_i, \bar{\partial}_i$:

$$dz_i(\partial_j) = d\bar{z}_i(\bar{\partial}_j) = \delta_{ij},$$

with the other values vanishing, $dz_i(\bar{\partial}_j) = d\bar{z}_i(\partial_j) = 0$ for all i, j .

Proposition 14.3. For a smooth function $f \in \mathcal{C}(U)$ its differential $df \in \Omega^1(U)$ can be split as the sum of $\Omega^{1,0}(U)$ and $\Omega^{0,1}(U)$ parts,

$$df = \partial f + \bar{\partial} f, \quad \partial f = \sum_1^n \partial_i f dz_i, \quad \bar{\partial} f = \sum_1^n \bar{\partial}_i f d\bar{z}_i.$$

This formula allows to extend uniquely the operator d to forms from $\Omega^{p,q}(U)$, placing the result into two parts, $\Omega^{p+1,0}(U)$ and $\Omega^{p,q+1}(U)$. The results are nontrivial for $p \leq n, q \leq n$.

14.9. Holomorphic forms and vector fields. A form ω is called holomorphic in U , if $\omega \in \Omega^{p,0}(U)$ and *in addition*, all coefficients of this form are holomorphic functions. Then the $(p, 1)$ -component of its differential is identically zero, $\bar{\partial}\omega = 0$.

A holomorphic vector field in U is a map picking an element of $T'_a M \simeq T_a M$ in a way holomorphically depending on a . In the same

way as in the real case, we can show that any holomorphic vector field has the form

$$X = \sum_{i=1}^n a_i(z) \frac{\partial}{\partial z_i}, \quad a_1, \dots, a_n \in \mathcal{O}(U).$$

Needless to say, holomorphic vector fields are closed by the Lie bracket.

15. Survey of further issues

15.1. Algebraic and analytic geometry. One can completely ignore the fact that $\mathbb{C} \supset \mathbb{R}$ and work only with subsets defined by polynomial equations with complex coefficients or (in the limit case) by converging series. The corresponding theory is very rich. Still in all what concerns the topology of the corresponding varieties, one is inevitably led to reintroducing the real subfield.

15.2. Integration. Another reason why one has to remember that $\mathbb{R} \subset \mathbb{C}$ is the fact that a complex (p, q) -form can be integrated only along submanifolds (chains) $N \subseteq M$ of *real* dimension $\dim_{\mathbb{R}} N = p + q$.

Example 15.1. For $n = 1$ a 1-form $\omega = a(z, \bar{z}) dz + b(z, \bar{z}) d\bar{z}$ can be integrated along curves in U . If $\omega \in \Omega^{1,0}(U)$, i.e., $b \equiv 0$ then $d\omega = \bar{\partial}\omega = \frac{\partial}{\partial \bar{z}} a dz \wedge d\bar{z}$. If ω is holomorphic, i.e., $\bar{\partial}a \equiv 0$, then $d\omega = 0$ and $\oint_{\gamma} \omega = 0$ for any curve γ homologous to zero in U .

The integral of $(1, 1)$ -form can be taken along 2-dimensional chains in ω .

15.3. Poincaré–Grothendieck theorem. Since $d = \partial + \bar{\partial}$, we have $0 = d^2 = \partial^2 + \partial\bar{\partial} + \bar{\partial}\partial + \bar{\partial}^2$. Applied to a form of the type (p, q) , this combination yields three terms of the type $(p + 2, q)$, $(p + 1, q + 1)$ and $(p, q + 2)$, respectively. Since the spaces $\Omega^{i,j}$ are disjoint, this implies the three separate identities,

$$\partial^2 = 0, \quad \partial\bar{\partial} + \bar{\partial}\partial = 0, \quad \bar{\partial}^2 = 0.$$

The (real) Poincaré theorem claims that if $U \subseteq \mathbb{R}^n$ is an open convex domain and $\omega \in \Omega^k(U)$, $k \geq 1$, such that $d\omega = 0$, then there exists $\alpha \in \Omega^{k-1}(U)$ such that $\omega = d\alpha$, that is, $\text{Ker } d = \text{Im } d$.

The analogous complex result, called the $\bar{\partial}$ -lemma or Grothendieck theorem, claims that for certain type of domains $U \in \mathbb{C}^n$ we have similar result, $\text{Ker } \bar{\partial} = \text{Im } \bar{\partial}$.

Theorem 15.1. *If $U = D_1 \times \cdots \times D_n$ is the product of one-dimensional domains $D_i \subseteq \mathbb{C}$ (possibly not simply connected) and $\omega \in \Omega^{p,q}(U)$ with $q \geq 1$ is such that $\partial\bar{\omega} = 0$, then there exists $\alpha \in \Omega^{p,q-1}(U)$ such that $\omega = \bar{\partial}\alpha$.*

15.4. Dolbeault cohomology. In a way mimicking the de Rham cohomology, we can define the Dolbeault cohomology as the quotient \mathbb{C} -spaces

$$H^{p,q}(M, \mathbb{C}) = \frac{\text{Ker } \bar{\partial} \subseteq \Omega^{p,q}(M)}{\text{Im } \bar{\partial} = \bar{\partial}(\Omega^{p,q-1}(M))}.$$

However, unlike the de Rham cohomology which is topologically invariant, the Dolbeault cohomology is not: it depends on the complex structure on the manifold M .

15.5. Hermitian and Kähler manifolds. The Hermitian inner product, as was noted, consists of the symmetric real part $\langle \cdot, \cdot \rangle$ and the antisymmetric imaginary part $[\cdot, \cdot]$.

Definition 15.1. A complex manifold is called *Hermitian*, if there is a Hermitian inner product defined on its *holomorphic* tangent bundle $TM = \bigcup_{a \in M} T_a M$, $\dim_{\mathbb{C}} T_a M = \dim M$.

The Hermitian structure h in local coordinates is defined by the tensor

$$\sum_{i,j} h_{ij}(z, \bar{z}) dz_i \otimes_{\mathbb{C}} d\bar{z}_j$$

such that the matrix $h_{ij}(z, \bar{z})$ is Hermitian positive definite.

The symmetric part $g = \frac{1}{2}(h + \bar{h})$ is a Riemannian metric on the holomorphic tangent bundle $T^{\mathbb{R}}M$ and it is symmetric bilinear on the corresponding complexification $T^{\mathbb{R}\mathbb{C}}M = \bigcup_a T_a^{\mathbb{R}\mathbb{C}}M$.

The imaginary part $\omega = \frac{i}{2}(h - \bar{h})$ is of type $(1, 1)$ on $T^{\mathbb{R}\mathbb{C}}M$ and is called the fundamental form. In local coordinates it takes the form

$$\sum_{i,j} \frac{i}{2} h_{ij} dz_i \wedge d\bar{z}_j.$$

Since $\bar{\omega} = \omega$, is real antisymmetric 2-form on on $T^{\mathbb{R}}M$.

Definition 15.2. A Hermitian manifold is called a *Kähler manifold*, if the fundamental form ω is closed, $d\omega = 0$ and hence defines a symplectic structure on M .

Example 15.2. The complex space \mathbb{C}^n with the standard “flat” Hermitian structure is a Kähler manifold. The projective space $\mathbb{C}P^n$ is Kähler with respect to a certain Hermitian structure called the *Fubini–Study metric*. The quotient tori also are.

Complex submanifolds of the above Kähler manifolds are hereditary Kähler, so plenty of examples can immediately be constructed.

Thus the Kähler manifolds possess simultaneously the three important structures (Riemann, symplectic and complex) that are in the natural agreement with each other. Needless to say, this richness brings a wealth of interesting results.

The equivalent Kählerianity conditions for dummies (here ∇ is the Riemannian covariant derivative associated with the Riemannian part of h):

$$\begin{aligned} d\omega &= 0 \\ \nabla\omega &= 0. \end{aligned}$$

15.6. Almost complex structure and Newlander–Nirenberg theorem. If $M^{\mathbb{R}}$ is the realification of a complex manifold and J an automorphism of the tangent bundle into itself, induced by the multiplication by the *constant* i , then the Lie bracket operation $X, Y \mapsto [X, Y]$ is bilinear over \mathbb{C} , so that $[JX, Y] = [X, JY] = J[X, Y]$.

Definition 15.3. An almost complex structure on an even-dimensional manifold is the operator field $J = \{J_a: T_aM \rightarrow T_aM\}$ (tensor) which a priori is variable (depending on a) satisfying the condition $J^2 = -\text{id}$.

All linear algebraic constructions can start not from a complex vector space V through its realification $V^{\mathbb{R}}$, but from the corresponding almost complex structure J on a real even-dimensional space. In particular, differential forms of the (p, q) -type can be defined for almost complex manifolds.

However, the derivation is in general looks different. In particular, for $\omega \in \Omega^{p,q}$ the exterior derivative $d\omega$ is a form of total degree $p+q+1$, but it may involve types other than just $(p+1, q)$ and $(p, q+1)$: all types (r, s) with $r+s = p+q+1$, $r, s \geq 0$, are possible.

For complex manifolds the above identity implies the integrability condition

$$[JX, JY] + [X, Y] - J[X, JY] - J[JX, Y] = 0.$$

It turns out that this condition is equivalent to the following conditions:

- (1) If X, Y are complex vector fields of type $(1, 0)$ with smooth coefficients, then $[X, Y]$ is also of type $(1, 0)$;
- (2) $d = \partial + \bar{\partial}$, that is, no “nonstandard” types appear by the exterior derivative,
- (3) $\bar{\partial}^2 = 0$.

Theorem 15.2 (A. Newlander–L. Nirenberg, 1957). *An integrable almost complex structure on an even-dimensional manifold M is induced*

by a genuine complex structure: there exists an atlas of complex charts on M with biholomorphic transition maps, such that $J = i$.

15.7. Global issues. Necessity of the sheaf theory. Unlike smooth functions, holomorphic functions are extremely rigid: a smallest “piece” (knowledge of the function in any open set) uniquely determines the possible continuation of this function with all singularities that may occur. In practical terms there is no partition of unity by holomorphic functions.

In practice this means that we cannot directly apply the “algebraic” formalism which derives geometric objects from the algebra $\mathcal{O}(M)$ of holomorphic functions. For instance, if M is a compact complex variety (e.g., a projective space), then there are no nonconstant functions holomorphic on it, so $\mathcal{O}(M) = \mathbb{C}$.

The alternative is to work with functions holomorphic only in parts of M . Yet such an approach makes it impossible to equip $\mathcal{O}(M)$ with algebraic operations: the sum of two functions is defined only if their domains have nonempty intersection, ditto other algebraic operations.

To overcome this seemingly “technical problem”, a special language should be developed. This language is known under the brand name *Sheaf Theory*. Alas, the time is over: to master this language, one needs another semester-long course.

References

- [N] R. Narasimhan, *Analysis on real and complex manifolds*, Advanced Studies in Pure Mathematics, Vol. 1, Masson & Cie, Éditeurs, Paris; North-Holland Publishing Co., Amsterdam, 1968. MR0251745
- [B] W. Ballmann, *Lectures on Kähler manifolds*, ESI Lectures in Mathematics and Physics, European Mathematical Society (EMS), Zürich, 2006. MR2243012

Additional literature

- [A] V. I. Arnold, *Ordinary differential equations*, Universitext, Springer-Verlag, Berlin, 2006. Translated from the Russian by Roger Cooke; Second printing of the 1992 edition. MR2242407
- [DFN1] B. A. Dubrovin, A. T. Fomenko, and S. P. Novikov, *Modern geometry—methods and applications. Part I*, 2nd ed., Graduate Texts in Mathematics, vol. 93, Springer-Verlag, New York, 1992. The geometry of surfaces, transformation groups, and fields; Translated from the Russian by Robert G. Burns. MR1138462
- [DFN2] ———, *Modern geometry—methods and applications. Part II*, Graduate Texts in Mathematics, vol. 104, Springer-Verlag, New York, 1985. The geometry and topology of manifolds; Translated from the Russian by Robert G. Burns. MR807945
- [KN1] S. Kobayashi and K. Nomizu, *Foundations of differential geometry. Vol. I*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1996. Reprint of the 1963 original; A Wiley-Interscience Publication. MR1393940
- [KN2] ———, *Foundations of differential geometry. Vol. II*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1996. Reprint of the 1969 original; A Wiley-Interscience Publication. MR1393941
- [M] S. Morita, *Geometry of differential forms*, Translations of Mathematical Monographs, vol. 201, American Mathematical Society, Providence, RI, 2001. Translated from the two-volume Japanese original (1997, 1998) by Teruko Nagase and Katsumi Nomizu; Iwanami Series in Modern Mathematics. MR1851352
- [G] C. Godbillon, *Géométrie différentielle et mécanique analytique*, Hermann, Paris, 1969 (French). MR0242081
- [C] I. Chavel, *Riemannian geometry*, 2nd ed., Cambridge Studies in Advanced Mathematics, vol. 98, Cambridge University Press, Cambridge, 2006. A modern introduction. MR2229062
- [J] J. Jost, *Riemannian geometry and geometric analysis*, 6th ed., Universitext, Springer, Heidelberg, 2011. MR2829653
- [B1] M. Berger, *A panoramic view of Riemannian geometry*, Springer-Verlag, Berlin, 2003. MR2002701
- [B2] W. M. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, 2nd ed., Pure and Applied Mathematics, vol. 120, Academic Press, Inc., Orlando, FL, 1986. MR861409

16. Exam

Problem 1. A tuple of smooth vector fields X_1, \dots, X_k on a smooth n -dimensional manifold M is *involutive*, if they are linear independent at each point $a \in M$ and the Lie bracket of any two vector fields belongs to the linear span of these fields, $[X_i, X_j] = \sum_k \alpha_{ijk} X_k$ for some functions $\alpha_{ijk} \in \mathcal{C}(M)$.

1. Prove that any other tuple Y_1, \dots, Y_k obtained from X_i by a nondegenerate linear transformation, $Y_i = \sum A_{ij} X_j$, $A_{ij} \in \mathcal{C}(M)$, with $\det \|A_{ij}(x)\|$ nowhere vanishing, is also involutive.

2. Give examples of involutive tuples.

A tuple of vector fields X_1, \dots, X_k is *integrable*, if for any point $a \in M$ there exists a piece of k -dimensional smooth submanifold N passing through a , such that all vector fields X_i are tangent to N . These submanifolds are called *integral submanifolds*.

3. Give examples of integrable tuples.

4. Prove that each integrable tuple is involutive.

5. Prove that a tuple which consists of a single nonvanishing vector field, is both integrable and involutive.

6. Assume that the tuple consists of *commuting* vector fields, $[X_i, X_j] = 0$ for all $i, j = 1, \dots, k$. Prove that it is involutive by explicitly constructing the integrable submanifolds using the flows of these fields.

7. Let X_1, \dots, X_k be an involutive tuple in \mathbb{R}^n locally near the origin. Show that there exists a system of local coordinates split into the two groups, "horizontal" (u_1, \dots, u_k) and "vertical" (v_{k+1}, \dots, v_n) , and a tuple of linear independent vector fields Y_1, \dots, Y_k , such that $Y_i = \frac{\partial}{\partial u_i} + V_i$, where V_i are tangent to the "vertical" direction and the tuples $\{X_i\}$ and $\{Y_i\}$ span the same subspace as in the item 1.

8. Prove that the tuple $\{Y_i\}$ is commutative as in item 5.

9. Prove that any involutive tuple of vector fields is integrable.

Problem 2. A differentiable 1-form $\omega \in \Omega^1(M)$ is called involutive, if $d\omega$ is divisible by ω , that is, there exists a 1-form $\eta \in \Omega^1(M)$ such that $d\omega = \omega \wedge \eta$.

1. Prove that an involutive form is integrable, that is, for any point $a \in M$ exists... (complete the statement).

2. Can you guess what should be the definition of involutivity for a tuple of 1-forms such that involutivity implies the integrability? Use the word "ideal".

Problem 3. The real 2-torus \mathbb{T}^2 is defined as the quotient space $\mathbb{R}^2/\mathbb{Z}^2$, hence the space $\mathcal{C}(\mathbb{T}^2)$ can be identified with the space of double periodic smooth functions on \mathbb{R}^2 .

1. From this description calculate the first and second de Rham cohomology of \mathbb{T}^2 .

The Möbius band can be defined as the quotient of the plane \mathbb{R}^2 by the identification $(x+1, -y) \simeq (x, y)$.

2. Calculate the de Rham cohomology of the Möbius band from this definition, including the accurate description of all relevant spaces (note that the Möbius band is non-compact).

Problem 4. Let \mathbb{D} be the unit disk $\{|z| < 1\} \subseteq \mathbb{C}$ with the Riemannian metric

$$\rho = \frac{2|dz|}{1-|z|^2}.$$

1. How do you understand this definition?

2. Show that the only *invertible* holomorphic map $f: \mathbb{D} \rightarrow \mathbb{D}$ which fixes the origin, is the rotation $f(z) = cz$, $|c| = 1$. *Hint.* This is the Schwarz lemma from Complex Analysis in disguise, which in turn follows from the maximum modulus principle applied to the holomorphic function $f(z)/z$. Show that the map

$$z \mapsto \frac{z-a}{1-\bar{a}z}, \quad a \in \mathbb{D},$$

is invertible holomorphic self-map of \mathbb{D} . We denote by G the group of holomorphic self-maps of \mathbb{D} . 3. Show that G consists of isometries of ρ : any $g \in G$ preserves ρ .

4. Show that each transformation from G is *circular*: it sends circles in \mathbb{C} to circles, if lines are considered as particular case of the circles.

5. Prove that the shortest path from any point $a \in \mathbb{D}$ to the origin is the segment of the corresponding radius. Prove that the diameters are geodesic curves for ρ .

6. Using symmetries from G , find and describe other geodesics, not passing through the origin. Describe all geodesics passing through a point $0 \neq a \in \mathbb{D}$. Prove that for any two points $a \neq b \in \mathbb{D}$ there exists at least one circle that passes through a and b and is orthogonal to the unit circle $\{|z| = 1\}$, called the *absolut*.

7. Let $\mathbb{H} = \{\text{Im } z > 0\} \subset \mathbb{C}$ be the upper half-plane and

$$\rho' = \frac{|dz|}{\text{Im } z}$$

a Riemannian metric on it. Show that \mathbb{H} and \mathbb{D} are biholomorphically equivalent, and each conformal map $f: \mathbb{D} \rightarrow \mathbb{H}$ is an isometry conjugating ρ' with ρ . Describe the geodesic curves on \mathbb{H} .

8. Show that the Gauss curvature of the metric ρ' is equal to -1 . *Hint.* The word *show* can be (only here) understood literally: if you can find these statements in a textbook, show me where.

9*. Suggest a Riemannian manifold with constant negative curvature $-\frac{1}{R^2}$, $1 \neq R$.

Problem 5. The complex projective line $\mathbb{C}P^1$ is formally defined as the quotient space $\mathbb{C}^2 \setminus \{0\}$ by $\mathbb{C} \setminus \{0\}$, that is, two points (x, y) and (u, v) are identified, if and only if there exists $\lambda \neq 0$ such that $(x, y) = \lambda(u, v)$. In the same way the projective plane $\mathbb{C}P^2$ is defined as the quotient $\mathbb{C}^3 \setminus \{0\} / \mathbb{C} \setminus \{0\}$.

1. Construct an atlas which makes $\mathbb{C}P^1$ into a holomorphic 1-dimensional manifold (holomorphic curve). Do the same for $\mathbb{C}P^2$. Show that $\mathbb{C}P^1 = \mathbb{C}^1 \sqcup \{*\}$, that is, the projective line can be obtained by adding a single point “at infinity” to the complex plane, which makes it topologically a sphere (*Riemann sphere*). Show $\mathbb{C}P^2 = \mathbb{C}^2 \sqcup \mathbb{C}P^1$, that is, the projective plane can be obtained by adding to the affine plane \mathbb{C}^2 the “line at infinity” $\mathbb{C}P^1$. *Hint.* The word *show* can be (only here) understood literally: if you can find these statements in a textbook, show me where.

2. Construct a nontrivial holomorphic vector field on $\mathbb{C}P^1$.

3. Construct a nonvanishing holomorphic vector field on $\mathbb{C}P^1$ or explain why it is impossible. *Hint:* what can be said about the sum total of the order of zeros and order of poles of any meromorphic function on $\mathbb{C}P^1$?

4. Construct a holomorphic 1-form on $\mathbb{C}P^1$ or explain why this is impossible.

5. Consider the circle $C = \{x^2 + y^2 = 1\} \subset \mathbb{R}^2$. Show that the differential 1-form $\omega = \frac{dx}{y}$ is smooth on C .

6. Construct a rational map $U: \mathbb{C}P^1 \rightarrow \mathbb{C}P^2$ such that $U|_{\mathbb{R}} \subset C$ (this map is called *rational uniformization*). Show how this map allows to find all rational points on the circle \mathbb{C} and all Pythagorean triples of natural numbers.

7. Compute the pullback $U^*\omega$ and show that it is meromorphic. Find its singular points and explain their location. *Hint.* A polynomial is holomorphic on \mathbb{C}^2 but has singularities on the infinite line after extension on $\mathbb{C}P^2$.

8. Consider the algebraic curve $E = \{y^2 = p(x)\} \subseteq \mathbb{R}^2$, where p is a real polynomial of degree 3 with simple roots. Prove that the 1-form

ω is holomorphic on the projective compactification of its complexification $\overline{E^{\mathbb{C}}} \subset \mathbb{C}P^2$. *Hint.* Where the troubles can come from? see the previous question.

9*. Find sufficient conditions guaranteeing that the 1-form $x^k \frac{dx}{y}$ is holomorphic on the *hyperelliptic curve* defined by the equation $\{y^2 = p(x)\} \subseteq \mathbb{C}P^2$ with $\deg p = n$.