Calculus on Manifolds

Sergei Yakovenko

Abstract. These are the lecture notes slightly revised and updated compared to the previous version of about a year ago.

They are highly informal and were aimed to be a supplement to more traditional expositions (see the list of recommended sources, mostly perennial classics). Lots of statements are on purpose formulated in a rather vague form stressing the principal features while sweeping under the carped technical and obscuring details (the devil is always in details, but can be ignored on the first date).

In essence, they introduce and motivate the algebra that is behind rather natural geometric constructions.

Date: 2017/8 academic year.

1. Introduction. Crash course on the Multivariate Calculus

1.1. **Linear algebra.** Field of real numbers $\mathbb{R}$, its origin (from $\mathbb{N}, \mathbb{Z}$ and $\mathbb{Q}$) and completeness. Real line $\mathbb{R}^1$ and real spaces $\mathbb{R}^n$. Linear functionals, linear maps. Compositions. Rank. Group $\mathrm{GL}(\mathrm{n}, \mathbb{R})$, determinant. Linear algebra as a universal paradigm. Matrix algebra vs. coefficients: vectors as single elements rather than collections of numbers; A as a "box" (operator) vs. $A = \big(a_{ij}\big)$, 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. Algebraic closure $\mathbb{C}$, its miraculous properties.

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

$$f\colon \mathbb{R}^m \to \mathbb{R}^n, \qquad f(x) = f(a) + A(x-a) + o(|x-a|), \tag{1}$$

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

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

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, $\mathrm{d}(f \circ g)(a) = \mathrm{d}f(b) \cdot \mathrm{d}g(a)$, where $b = g(a)$.

Remark 1.1. Existence of differentiable maps (functions) is a *non-trivial* circumstance which is often obscured by the fact that all polynomial, rational and algebraic maps are differentiable. This creates a false impression that "almost all" functions are differentiable, which is plain wrong. Most dynamic processes in our world, with a few happy exceptions in celestial mechanics, are non-differentiable.

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_aU; it is isomorphic to $\mathbb{R}^n$.

The union $TU = \bigcup_{a \in U} T_aU$ 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_aU. In the same vein the image Av is identified with the vector tangent to T_bV, $b = f(a)$. This is the convention that will be always in place.

Definition 1.4. Diffeomorphism $f\colon U \to V$ for $U, V \subset \mathbb{R}^n$ is a one-to-one smooth map whose set theoretic inverse $f^{-1}\colon V \to U$ 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, than 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 $\psi \mapsto -h \circ (\mathrm{id} + \psi)$ 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. □

1.4. **Dimension one: a notable exception.** The strategic difference between functions and maps comes from the fact that $\mathbb{R}$ has the structure of a field and also that 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 $f\colon \mathbb{R}^1 \to \mathbb{R}^1$ 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 $f'\colon a \mapsto f'(a)$ 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 considerall real-valued *smooth* (= infinitely differentiable) functions in U, which constitute an $\mathbb{R}$-algebra, i.e., a linear (infinite-dimensional) topological space $\mathscr{C}(U)$ over the field $\mathbb{R}$ with the multiplication operation (sometimes we will use the more verbose notation $C^\infty(U)$ instead of $\mathscr{C}(U)$ or just $\mathscr{C}$).

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 1.2. For a smooth function its differential $\mathrm{d}f$ at a point $a \in U$ is a linear map from $T_aU \simeq \mathbb{R}^n$ 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 $\gamma\colon U \to V \subseteq \mathbb{R}^n$, 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 $\mathrm{d}\gamma(a)\colon T_aU \to T_{\gamma(a)}V$. The image $v(a) = \mathrm{d}\gamma(a)\mathbf{1}$ 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 $\dot\gamma(a) = v \in T_aU$, we say that the curve γ is tangent to the vector v at the point a.

2. Vector fields on $\mathbb{R}^n$ 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 $X\colon U \to TU$ such that $X(a) \in T_aU$ 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 $\mathscr{X}(U)$.

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 $X \in \mathscr{X}(U)$ is a smooth vector field, then for any point $a \in U$ there exists a unique smooth curve $\gamma_a\colon (\mathbb{R}, 0) \to (U, a)$ 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 $\widetilde\gamma_a$ 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 $\gamma\colon t \mapsto x(t)$ 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

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = v_i(x_1(t), \dots, x_n(t)), \quad x(0) = a, \qquad i = 1, \dots, n.$$

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

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

(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 $\frac{\mathrm{d}x}{dt} = x^2$ 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 $t \leqslant 0$. 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 $X \in \mathscr{X}(U)$ by the field $\widetilde{X} = \lambda X$, $\lambda \in \mathbb{R}_{\text{field}}$, then one can easily integrate the new equation by setting $\widetilde{\gamma}_a(t) = \gamma_a(\lambda t)$. In the more general case $\widetilde{X} = \varphi X$, $\varphi \in \mathscr{C}(U)$, one can also find solutions by an integration. In particular, if $\varphi > 0$, then the images of the curves γ_a and $\widetilde{\gamma}_a$ *locally* coincide.

2.2. **Flows of vector fields.** One can consider simultaneously all integral curves of a vector field $X \in \mathscr{X}(U)$ 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 $F_X \colon (t, a) \mapsto F_X^t(a)$ 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. □

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

$$F^t \circ F^s = F^{s+t} = F^s \circ F^t, \qquad F^0 = \mathrm{id}.$$

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

The map $F_X \colon \mathbb{R} \times U \to U$ 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 nondegenerate (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* $\mathscr{C}(U)$ 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). This algebra will be referred to as the *structural algebra* of U.

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

The correspondence $F \mapsto F^*$ is *functorial*, in particular, if $F: U \to V$ and $G: V \to W$, so that $G \circ F: U \to W$, then $(G \circ F)^*: \mathscr{C}(W) \to \mathscr{C}(U)$ 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 $A: \mathscr{C}(V) \to \mathscr{C}(U)$ is induced by some smooth map F, so that $A = F^*$.

Proof. Consider the coordinate functions $y_i \in \mathscr{C}(V)$ and look at the map $F = (f_1, \dots, f_n)$, where $f_i = Ay_i$. By homomorphy, $Au = F^*u$ 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 $X \in \mathscr{X}(U)$ is a smooth vector field, then any function $\varphi: U \to \mathbb{R}$ can be differentiated along X producing a new function, denoted sometimes $X\varphi$, sometimes $L_X \varphi \in \mathscr{C}(U)$.

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

$$X\varphi(a) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \varphi(F_X^t(a))$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \varphi(\gamma(t)) = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i}(a) \cdot \frac{\mathrm{d}x_i(t)}{\mathrm{d}t}\bigg|_{t=0} = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i}(a)\, v_i(a).$$

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 $X: \mathscr{C}(U) \to \mathscr{C}(U)$, is an $\mathbb{R}$-linear operator which satisfies the *Leibniz rule*,

$$X(\varphi\psi) = \varphi \cdot (X\psi) + \psi \cdot (X\varphi). \tag{2}$$

Conversely, if $L: \mathscr{C} \to \mathscr{C}$ is an $\mathbb{R}$-linear self-map of $\mathscr{C}$ 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, $v_i = Lx_i$, $x_i \in \mathscr{C}$. 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

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

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

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

Lie derivations (first order differential operators on $\mathscr{C} = \mathscr{C}(U)$, satisfying the Leibniz rule) form a *module* over the algebra $\mathscr{C}$: they can be added between themselves and multiplied by functions from $\mathscr{C}$. 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 $\mathscr{X}(U)$ take the algebra $\mathscr{C}(U)$ 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 $X, Y \in \mathscr{X}(U)$ and $\varphi, \psi \in \mathscr{C}(U)$, 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 $X, Y \in \mathscr{X}(U)$ 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 $X \in \mathscr{X}(U)$ and the respective flow $\{F_X^t \in \operatorname{Aut}(U) : t \in \mathbb{R}\}$ is expressed by the formula

$$\left.\frac{\mathrm{d}F_X^t(a)}{\mathrm{d}t}\right|_{t=0} = X(a) \in T_aU,$$

the left hand side being the 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 $\mathscr{C}(U)$. Then for any test function $\varphi \in \mathscr{C}(U)$ we have a one-parametric family of functions $\{\varphi_t = (F_X^t)^*\varphi : t \in \mathbb{R}\}$ and can compute the (partial) derivative $\frac{\mathrm{d}\varphi_t}{\mathrm{d}t}\big|_{t=0}$. The answer is not surprising: the derivative is equal to Xf. In other words, we have the identity between the linear operators on $\mathscr{C}(U)$,

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

If the algebra $\mathscr{C}(U)$ were finite-dimensional, we would have a one-parametric group of linear operators $\{B^t\}$ such that $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} = A$ 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:

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

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 $\mathscr{C}(U)$.

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

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

It holds for all analytic functions $\varphi \in \mathscr{C}(U)$, 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 $F\colon U \to V$ be a smooth map. How it acts on vector fields?

Geometric push-forward. Suppose $X \in \mathscr{X}(U)$ is a vector field is on U. Take an arbitrary point $a \in U$ and consider a curve $\gamma_a\colon (\mathbb{R}^1, 0) \to (U, a)$ 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 $w = \mathrm{d}F(a) \cdot v$.

Algebraic push-forward. We need to construct a Lie derivation of $\mathscr{C}(V)$ which becomes the Lie derivative X on $\mathscr{C}(U)$ after the pull-back F^*. This is straightforward: for any $\varphi \in \mathscr{C}(V)$ consider $\psi = F^*\varphi$ and apply X. The resulting operator $\mathscr{C}(V) \to \mathscr{C}(U)$ will satisfy the Leibniz rule as required, but we need to get back to $\mathscr{C}(V)$. In general it is impossible, but if F is a diffeomorphism, then we can use its inverse F^{-1} to to the job and define $Y\colon \mathscr{C}(V) \to \mathscr{C}(V)$, $Y = (F^{-1})^* X F^*$. 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 $\mathscr{C}(V) \to \mathscr{C}(U)$. 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 $F\colon \mathbb{R}^2 \to \mathbb{R}^1$ 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 $\mathscr{C}(U)$, 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 $\mathscr{C}(U)$ 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 $X \in \mathscr{X}(U)$ 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 $\chi \in \mathscr{C}(U)$ 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 $x\colon U \to \mathbb{R}^n$ 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 $x\colon U \to \mathbb{R}^n$ and $y\colon V \to \mathbb{R}^n$ defined on the overlapping domains $U \cap V \neq \varnothing$ are said to agree on the intersection, if the *transition map* $y \circ x^{-1}\colon x(U \cap V) \to \mathbb{R}^n$ 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.

3.3. **Chart-wise approach to working on manifolds.** Assume that we are given a smooth manifold with an atlas of charts $x_\alpha\colon U_\alpha \to \mathbb{R}^n$, $\alpha \in \mathfrak{A}$ (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 3.2. A function $f\colon M \to \mathbb{R}$ is called C^∞-smooth (or simply smooth) on M, if all compositions $f_\alpha\colon f \circ x_\alpha^{-1}$, $\alpha \in \mathfrak{A}$, are C^∞-smooth on the domains $D_\alpha = x_\alpha(U_\alpha) \subseteq \mathbb{R}^n$.

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

Any function $f \in \mathscr{C}(M)$ 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 $D_{\alpha\beta} = x_a(U_\alpha \cap U_\beta)$ and $D_{\beta\alpha} = x_\beta(U_\alpha \cap U_\beta)$:

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

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

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

The set of all tangent vectors at a point a is denoted by T_aM.

Each chart x_α provides an isomorphism between T_aM, $a \in M$, 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

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

Definition 3.4. A vector field on a manifold M is a smooth map $X\colon M \to TM$ such that $X(a) \in T_aM$.

One can associate with any $X \in \mathscr{X}(M)$ 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).

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

Definition 3.5. A map $\gamma\colon (\mathbb{R}, 0) \to M$ 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 3.6. 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 th structure of a linear (vector) space. This can be proved by computation, OK, but remember about of "theft vs. toil".

4. Constructions on manifolds

4.1. Points, curves, functions.

Definition 4.1.

Manifold: a Hausdorff second-countable (not too large) topological space, equipped with a self-consistent atlas of local charts.

Local chart: a homeomorphism $x\colon U \to B$ between an open subspace $U \subseteq M$ and an open ball $B \subseteq \mathbb{R}^n$.

Self-consistent atlas: collection of charts $\{x_\alpha\colon U_\alpha \to B_\alpha\}$, such that all transition maps $F_{\alpha\beta} = x_\alpha \circ x_\beta^{-1}$ are diffeomorphisms in their natural domains $B_{\alpha\beta} = x_\beta(U_{\alpha\beta})$, $U_\alpha\beta = U_\alpha \cap U_\beta$.

Example 4.1. Euclidean spaces of all dimensions are smooth manifolds with atlases consisting of single chart each.

Definition 4.2. A continuous map $F\colon M \to N$ between two manifolds is smooth if for any local charts x_α on M and y_β on N the composition $y_\beta \circ F \circ x_\alpha^{-1}$ is C^∞-smooth as a map between two balls of the respective Euclidean spaces.

A smooth function is a continuous function $f\colon M \to \mathbb{R}$ such that the composition $f \circ x_\alpha^{-1}$ is smooth in each chart x_α.

A smooth curve on M is a map $\gamma\colon \mathbb{R}^1 \to M$ such that $x_\alpha \circ \gamma$ is smooth.

Example 4.2. If we fix a local chart, then $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, can be used to denote a point in M (with the corresponding coordinates in B). Then smooth function becomes an explicit expression of n real variables $f(x) = f(x_1, \dots, x_n)$. In the same way a smooth curve is defined by a smooth vector function $\gamma = x(t) = (x_1(t), \dots, x_n(t))$.

4.2. Structural algebra.

Definition 4.3. The commutative algebra $\mathscr{C} = \mathscr{C}_M = C^\infty(M)$ of smooth (scalar) functions on M is called the *structural algebra* of M.

Definition 4.4. For any smooth map $F\colon M \to N$ the *pullback* of F if the map between the structural algebras

$$F^*\colon \mathscr{C}_N \to \mathscr{C}_M, \qquad g \mapsto f = F^*g = g \circ F, \quad f(x) = g(F(x)).$$

Note the inversion of the direction.

Proposition 4.1. *The map F^* is a morphism between the algebras: it is an $\mathbb{R}$-linear maps which preserves the multiplication:*

$$F^*(g_1 g_2) = F^*g_1 \cdot F^*g_2.$$

In the linear algebra it is customary to omit braces around the argument and write Ax rather than $A(x)$ to denote the value taken by A on the vector x. We shall stick to this notation.

4.3. **Tangent vectors.** Let $a \in M$ be a point.

Definition 4.5. Two curves $\gamma, \gamma' \colon (\mathbb{R}^1, 0) \to M$ are *tangent to each other* at a, if $\gamma(0) = \gamma'(0) = a$, and in any chart x we have

$$(x \circ \gamma)(t) - (x \circ \gamma')(t) = o(t) \qquad \text{as } t \to 0.$$

Remark 4.1. In general, points of the manifold cannot be added (or subtracted) between themselves. The above definition uses the fact that the subtraction is well defined in $\mathbb{R}^n$.

One can instantly see that this definition is independent of the choice of the chart x and is indeed an equivalence relation (reflexive, transitive and symmetric).

Definition 4.6. The *tangent space* T_aM at a point $a \in M$ is the set of equivalence classes of smooth curves through a.

The *tangent bundle* is the disjoint union of all tangent spaces,

$$TM = \bigsqcup_{a \in M} T_aM.$$

Remark 4.2. In coordinates any smooth curve is represented by the parametric expression $\gamma(t) = (x_1(t), \dots, x_n(t))$ with $\gamma(0) = x(a)$, so that $x(t) = a + tv + o(t)$, $v = (v_1, \dots, v_n) \in \mathbb{R}^n$.

However, the above definition defines T_aM only as a point set in one-to-one correspondence with $\mathbb{R}^n$ (the correspondence depends on the choice of the chart). No structure of a vector space is defined yet.

Definition 4.7. Let $f \in \mathscr{C}_M$ and $v \in T_aM$ a tangent vector represented by a curve $\gamma \colon (\mathbb{R}, 0) \to M$, $\gamma(0) = a$.

The *directional derivative* is the number

$$\frac{\mathrm{d}f}{\mathrm{d}v}(a) = \left.\frac{\mathrm{d}(f \circ \gamma)(t)}{\mathrm{d}t}\right|_{t=0}.$$

Using the chain rule, one can easily check that this definition does not depend on the choice of the curve γ representing v.

The operator

$$\frac{\mathrm{d}}{\mathrm{d}v} \colon \mathscr{C}_M \to \mathbb{R}, \qquad f \longmapsto \frac{\mathrm{d}f}{\mathrm{d}v}(a)$$

is $\mathbb{R}$-linear and satisfies the Leibniz rule,

$$\frac{\mathrm{d}(fg)}{\mathrm{d}v}(a) = f(a) \cdot \frac{\mathrm{d}g}{\mathrm{d}v}(a) + g(a) \cdot \frac{\mathrm{d}f}{\mathrm{d}v}(a)$$

This property can be used as a definition of the directional derivative at a point a.

Definition 4.8. Let $a \in M$. A(n abstract) *directional derivative* is an operator $X_a\colon \mathscr{C}_M \to \mathbb{R}$ which is $\mathbb{R}$-linear and satisfies the Leibniz rule,

$$X_a(f+g) = X_a f + X_a g, \qquad X_a(fg) = f(a)\,X_a(g) + g(a)\,X_a(f).$$

Proposition 4.2. *Directional derivatives form a vector space over $\mathbb{R}$: the sum of directional derivatives again satisfies the above conditions (ditto λX).*

Remark 4.3. If $M = \mathbb{R}^n$, then one can prove (using the Hadamard lemma) that any abstract directional derivative is a genuine directional derivative along a certain vector.

Remark 4.4. This proposition allows to equip each T_aM with the structure of a $\mathbb{R}$-vector space, which was difficult to do before.

4.4. **Push-forward of tangent vectors.** If $F\colon M \to N$ is a smooth map and $a \in M$, then one can define the push-forward

$$(F_*)_a\colon T_aM \to T_bN, \qquad b = F(a),$$

by two equivalent ways.

Definition 4.9 (Geometric definition)**.** For a tangent vector $v \in T_aM$ represented by a smooth curve $\gamma\colon (\mathbb{R}^1, 0) \to (M, a)$, the push-forward $(F_*)_a v$ is the directional derivative along the smooth curve

$$F \circ \gamma\colon (\mathbb{R}^1, 0) \to (N, b), \qquad b = F(a).$$

Of course, independence of the choice of the representative γ has to be verified.

Definition 4.10 (Algebraic definition)**.** Define the operator

$$Y_b\colon \mathscr{C}_N \to \mathbb{R}, \qquad Y_b = X_a \circ F^*.$$

Since F^* is a morphism of algebras and X_a satisfies the Leibniz rule, Y_b also does.

A trivial exercise is to show that the two definitions are equivalent.

5. Vector fields

5.1. **Basic definitions.** Intuitively a vector field on a manifold M is a collection of tangent vector $v(a) \in T_aM$ attached to each point a, which depends smoothly on a. The smoothness condition means that we deal with a smooth map $v\colon M \to TM$ such that for any $a \in M$ we have $v(a) \in T_aM$. But then again the problem reduces to equipping the set $TM = \bigsqcup_{a\in M} T_aM$ with the structure of a smooth manifold.

It is an extremely useful exercise to transform any atlas of smooth charts on M to an atlas of smooth charts on TM which would make it into a smooth manifold of dimension $2n$. Development of this idea leads to the very useful notion of a general *vector bundle* over a manifold.

The algebraic approach is much more straightforward.

Definition 5.1. A smooth vector field on M is a derivation of the structure algebra $\mathscr{C}_M$, i.e., an $\mathbb{R}$-linear operator X satisfying the Leibniz rule,

$$X\colon \mathscr{C}_M \to \mathscr{C}_M, \qquad X(f\cdot g) = f\cdot Xg + g\cdot Xf.$$

Indeed, evaluation X_a at an arbitrary point $a \in M$ will be an abstract directional derivative at a.

The set of smooth vector fields on M will be denoted by $\mathscr{X}(M)$, $\mathscr{X}_M$ or simply $\mathscr{X}$ when M is clear from the context.

If $F\colon M \to N$ is a smooth map, then **in general F does not allow to carry vector fields in any direction**. There is an exception, however: if F is a diffeomorphism, then it allows both push forward and pull-back of vector fields.

Definition 5.2. Let $F\colon M \to N$ be a smooth map, $X \in \mathscr{X}(M)$, $Y \in \mathscr{X}_N$. We say that X and Y are F-related, if $F^*Y = XF^*$, i.e., if the following diagram is commutative,

$$\begin{tikzcd}
\mathscr{C}_M \arrow[d, "X"] & \mathscr{C}_N \arrow[l, "F^*"'] \arrow[d, "Y"] \\
\mathscr{C}_M & \mathscr{C}_N \arrow[l, "F^*"']
\end{tikzcd}$$

In particular, if F is a diffeomorphism, then $Y = (F^*)^{-1}XF^* \in \mathscr{X}_N$ is a push-forward of $X \in \mathscr{X}_M$, and, inversely, $X = F^*Y(F^*)^{-1}$ is a pullback of X.

In other words, X, Y are similar (conjugated by an isomorphism F^* between the structural algebras). Sometimes we will use the notation

$$F_*\colon \mathscr{X}_M \to \mathscr{X}_N, \qquad F_*X = (F^*)^{-1}XF^*,$$
$$F^*\colon \mathscr{X}_N \to \mathscr{X}_M, \qquad F^*Y = F^*Y(F^*)^{-1}.$$

Remark 5.1 (important)**.** For invertible F both push-forward and pull-back operators enjoy equal "legal status" and can be applied. However, there is a subtle reason to prefer the pull-back.

The (infinite-dimensional) $\mathbb{R}$-vector space $\mathscr{X}_M$ is naturally a $\mathscr{C}_M$-module over the structural algebra of smooth functions: any field X can be multiplied by a function f producing the field fX. Since for the structural algebra the pull-back is much more natural operation, so should be for the fields. Then

$$F^*(fX) = F^*f \cdot F^*X,$$

which makes the pull-back into the morphism of modules.

5.2. **Flow of vector fields.** Let $v \in \mathscr{X} = \mathscr{X}(M)$ be a vector field on a manifold M.

Definition 5.3. The flow of v is a smooth map $\mathbb{R} \times M \to M$, denoted by $(t, x) \mapsto F^t(x)$, such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} F^t(x) = v(x) \qquad \forall t \in \mathbb{R}, \quad x \in M.$$

If it is necessary to indicate the field, we will write $F^t_v(x)$.

Theorem 5.1 (local existence theorem for flows)**.** *For any vector field v and any point $a \in M$ there exists a neighborhood $(\mathbb{R}, 0) \times (M, a)$ such that the flow map as above is defined in this neighborhood,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} F^t(x) = v(x) \qquad \forall t \in (\mathbb{R}, 0), \quad x \in (M, a). \quad \square$$

Remark 5.2. In general, the flow can indeed be defined only locally on open neighborhoods $(\mathbb{R}, 0) \times (M, a)$. It is globally defined if M is compact.

To avoid (mostly, technical) difficulties and simplify the exposition, we will assume that all flows are globally defined.

Proposition 5.1. *Considered as a family of self-maps*

$$\{F^t \colon M \to M\}_{t\in\mathbb{R}},$$

the flow forms a one-parametric commutative group isomorphic to $\mathbb{R}$*,*

$$F^{t+s} = F^t \circ F^s = F^s \circ F^t \qquad \forall t, s \in \mathbb{R}.$$

In particular, all maps are invertible (diffeomorphisms),

$$(F^t)^{-1} = F^{-t}, \qquad F^0 = \mathrm{id}\,.$$

The algebraic construction seems to be more succinct. Note that the family of automorphisms $\{(F^t)^*\}_{t\in\mathbb{R}}$ of the structural algebra $\mathscr{C}_M$ will also satisfy the group property.

Definition 5.4. A one-parametric group $\{A^t\colon \mathscr{C}_M \to \mathscr{C}_M\}_{t\in\mathbb{R}}$ of automorphisms of the structural algebra is said to be *generated* by an operator X, if

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} A^t = X,$$

Applying this operator identity to a test function f at a point $a \in M$, we see that it is simply the definition of the directional derivative of f for the special curve $\gamma(t) = F^t(a)$.

It is an elementary calculation to show that X is a derivation, $X \in \mathscr{X}_M$. Indeed, by the Leibniz rule

$$\begin{aligned} X(fg) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \big(A^t(fg)\big) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \big(A^t f \cdot A^t g\big) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} A^t f \cdot A^0 g + A^0 f \cdot \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} A^t g = X(f)g + fX(g). \end{aligned}$$

Remark 5.3. The problem of reconstruction of the family $\{A^t\}$ from the derivation $X \in \mathscr{X}_M$ is much more subtle: it is equivalent to the global integration of ODEs (vector fields) and the result depends on M. Formally one can use the exponential series and define

$$A^t = \mathrm{id} + X + \frac{1}{2!}X^2 + \cdots + \frac{1}{k!}X^k + \cdots, \qquad t \in \mathbb{R},$$

but the convergence can be justified only for finite-dimensional algebras (or on dense subsets of $\mathscr{C}_M$). We will simply use the group of automorphisms $A^t = (F^t)^*$ of the structural algebra.

Still this similarity is suggestive enough that some sources denote the one-parameter group of isomorphisms of the structural algebra, using the exponential notation,

$$A^t = \exp tX \qquad \text{or} \qquad A^t = \mathrm{e}^{tX}.$$

6. Lie derivative: yet another object to derive

6.1. **Lie derivative of vector fields.** Assume now that we have *two* vector fields X, Y on the same manifold. Can one define the "directional derivative" of Y along X in the same way as we defined it for functions? *Voilá.*

Definition 6.1. Denote by $(F_X^t)^*\colon \mathscr{X}_M \to \mathscr{X}_M$ the family of pull-back operators associated with the flow of the vector field X.

We define the directional derivative of Y along X in the most natural way,

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

The two notations, the Lie notation L_X and the pseudoclassical notation $\frac{\mathrm{d}Y}{\mathrm{d}X}$ will be used in the fully interchangeable way.

By construction, Z is an operator from $\mathscr{C}_M$ to itself which satisfies the Leibniz rule. It can be easily computed.

$$\begin{aligned}\frac{\mathrm{d}Y}{\mathrm{d}X} f &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (F_X^t)^* Y f = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Big((F^t)^* Y (F^{-t})^*\Big) f \\ &= \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (F^t)^* Y - Y \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (F^t)^*\right) f = (XY - YX) f.\end{aligned}$$

Definition 6.2. The expression $XY - YX$ is called the *commutator*, or the *Lie bracket* of vector fields $X, Y \in \mathscr{X}$, and denoted by

$$[X, Y] = XY - YX.$$

By construction, the commutator is again a vector field, $[X, Y] \in \mathscr{X}$.

Remark 6.1. The above identity can be restated as follows,

$$L_{[X,Y]} = [L_X, L_Y], \qquad L_X, L_Y, L_{[X,Y]} \quad \text{Lie derivations of } \mathscr{C}.$$

Remark 6.2. This equality is somewhat surprising. First, it is totally unexpected that

$$\frac{\mathrm{d}X}{\mathrm{d}Y} = -\frac{\mathrm{d}Y}{\mathrm{d}X},$$

that is, the role of the two fields is almost symmetric.

Second surprising fact is that the commutator is again a first order operator: when computing its action on a function, we observe that the second order derivatives cancel each other and only the first partial derivatives remain.

This is less surprising if we recall that the Lie derivative of any object is by construction the object of the same type (we shall see more examples of this principle)

6.2. **Properties of the commutator.** The Lie bracket

$$[\,\cdot\,,\,\cdot\,] \colon \mathscr{X} \times \mathscr{X}$$

is a $\mathbb{R}$-bilinear form which is Leibniz in each argument:

$$[X, fY] = [X, f] \cdot Y + f \cdot [X, Y],$$

where we used the notation $[X, f]$ for the Lie derivative of f along X, which is of course the same as the usual Xf.

If $Z \in \mathscr{X}$ is a third vector field and

$$\tfrac{\mathrm{d}}{\mathrm{d}Z}\colon \mathscr{X} \to \mathscr{X}, \qquad \tfrac{\mathrm{d}}{\mathrm{d}Z}X = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (F_Z^t)^* X,$$

the corresponding Lie derivative, then because of the above bilinearity we have the Leibniz rule,

$$\tfrac{\mathrm{d}}{\mathrm{d}Z}[X, Y] = [\tfrac{\mathrm{d}}{\mathrm{d}Z}X, Y] + [X, \tfrac{\mathrm{d}}{\mathrm{d}Z}Y].$$

Expressing the Lie derivative as a commutator, we see that

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

This is the *Jacobi identity*, which is usually formulated (and more easily memorized) as vanishing of the cyclical sum

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

The "standard" proof of this identity is by replacing each double commutator with four third order operators involving composition of X, Y, Z in all possible orders. Altogether we get 12 combinations which exhaust 6 possible permutations, each occurring twice with opposite signs. Of course, it is impossible to recognize the Leibniz rule in this combinatorial rampage.

Remark 6.3. One can also easily check that the identity

$$[L_X, L_Y] = L_{[X,Y]}, \qquad \text{where} \quad [L_X, L_Y] = L_X L_Y - L_Y L_X,$$

valid for the action of Lie derivatives on $\mathscr{C}$, holds also if we consider the action of both sides on $\mathscr{X}$. This is yet another reincarnation of the Jacobi identity.

In coordinates, if we denote the Jacobian matrixes of vector fields $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots, w_n)$ by $\left(\frac{\partial v}{\partial x}\right)$ and $\left(\frac{\partial w}{\partial x}\right)$ respectively, then the commutator of $X = \sum v_i(x)\frac{\partial}{\partial x_i}$ and $Y = \sum w_i(x)\frac{\partial}{\partial x_i}$ will be the field with the coordinates

$$\left(\frac{\partial v}{\partial x}\right) w - \left(\frac{\partial w}{\partial x}\right) v.$$

Remark 6.4 (warning). In the above calculations we used several times the notation

$$L_v = \frac{\mathrm{d}}{\mathrm{d}v} \quad \text{or} \quad L_X = \frac{\mathrm{d}}{\mathrm{d}X}$$

for derivative along the vector field. This allowed to save space, but in general this may result in confusion.

If $f \in \mathscr{C}$ is a smooth function, how the two operators, L_{fv} and L_v are related? you might expect one of the two things, either "tensiorial" behavior whereby $L_{fv} = fL_v$, or the "Leibniz-type" behavior, whereby $L_{fv} = f \cdot L_v + (L_v f)\,\cdot$ (the second term acts as the multiplication by the function $g = L_v f$).

In reality the answer depends on whether L_v is considered on functions (and then the first case occurs) or on vector fields (and then the second possibility is realized). Later we will describe how L_v acts on exterior forms, where the answer will take yet another form.

In short, be careful with the Lie derivative!

6.3. **Commutator and the flows.** If $[X, Y] = 0$, then this means that the flow F_X^t preserves the field Y and, reciprocally, F_Y^s preserves X. This means that each of the flows maps a trajectory of the other vector field to another such trajectory. This means that the two flows commute,

$$F_X^t \circ F_Y^s = F_Y^s \circ F_X^t, \qquad \forall t, s \in \mathbb{R}.$$

Conversely, if two flows are commuting in the above sense, then their velocity fields have identically zero commutator.

This observation paves the way to a very useful observation.

Two commuting flows have a common 2-dimensional integral surface: if at some point the vectors $X(a), Y(a)$ are linear independent, the map

$$(\mathbb{R}^2, 0) \to M, \qquad (t, s) \mapsto F_X^t \circ F_Y^s(a)$$

is locally injective and both fields are tangent to its image $S_a \subseteq M$. The variables (t, s) can be seen as the local coordinates on S_a, centered at the point a.

The mutual position of surfaces S_a for different points a can be understood from the following rectification theorem.

Theorem 6.1 (simultaneous rectification theorem)**.** *If $X, Y \in \mathscr{X}(\mathbb{R}^2, 0)$ are two commuting vector fields linear independent at the origin, then there exists a local diffeomorphism which simultaneously rectifies both fields into two constant fields.*

Proof. Assume that the first vector field X is already rectified and in the new coordinates it has the form $\frac{\partial}{\partial x_1}$. Commutation with this vector field means that

$$Y = \sum_{i=1}^{n} v_i(x)\,\frac{\partial}{\partial x_i}, \qquad \frac{\partial v_i}{\partial x_1} \equiv 0.$$

The linear independence condition implies that the vector $\sum_{i=2}^n v_i(0)\frac{\partial}{\partial x_i}$ is nonzero. Consider the diffeomorphism of the space $(\mathbb{R}^{n-1}, 0)$ equipped with the coordinates $(x_2, \dots, x_n)$ which rectifies the field

$$Y' = \sum_{i=2}^{n} v_i(x_2, \dots, x_n) \frac{\partial}{\partial x_i} \in \mathscr{X}(\mathbb{R}^{n-1}, 0).$$

Extended by the identical transformation of x_1-coordinate, this diffeomorphism preserves X and rectifies Y. $\square$

Corollary 6.1. *Common integral surfaces of two commuting vector fields locally look like a family of parallel* 2*-planes in* $\mathbb{R}^n$*.*

Existence of common integral surfaces is called *integrability* of the pair of fields. The above result states that tuples of pairwise commuting vector fields are integrable.

Remark 6.5. Of course, the result is true for any number of $k \leqslant n$ pairwise commuting vector fields.

6.4. **Frobenius theorem.** The condition of commutativity is not necessary for integrability: if two vector fields X, Y are tangent to the same surface $S \subset M$, then one can only claim that $[X, Y]$ is again tangent to M (why?), that is, $[X, Y]$ is in the submodule $\mathscr{C}_M(X, Y)$ of the module $\mathscr{X}(M)$ generated by X and Y:

$$[X, Y] = fX + gY, \qquad f, g \in \mathscr{C}(M).$$

Definition 6.3. A tuple of vector fields $X_1, \dots, X_k \in \mathscr{X}(M)$ is called *involutive*, if the submodule $\mathscr{C}_M(X_1, \dots, X_k)$ generated by them in the $\mathscr{C}(M)$-module $\mathscr{X}(M)$ is closed by the Lie bracket.

Of course, if X_i are commuting, then the Lie bracket is zero on the pairs of generators and preserves the module. Let $M = (\mathbb{R}^n, 0)$ and assume that $X_1, \dots, X_k \in \mathscr{X}_M$ are linear independent at the origin.

Theorem 6.2 (Frobenius)**.** *If* $\mathscr{C}_M(X_1, \dots, X_k)$ *is a submodule closed by the bracket, then locally one can find a system of commuting generators for the module: there exist* $Y_1, \dots, Y_k \in \mathscr{C}_M$ *such that*

$$\mathscr{C}_M(X_1, \dots, X_k) = \mathscr{C}_M(Y_1, \dots, Y_k), \qquad [Y_i, Y_j] = 0.$$

Proof. We start by choosing a convenient coordinate system. Assume that the first k coordinates, called "horizonal", are chosen in such a way that $X_1(0), \dots, X_k(0)$ are the basic vectors. Denote by $E_1, \dots, E_n$ the coordinate vector fields (commuting between themselves). By assumption,

$$X_i = E_i + \sum_{j=1}^{n} a_{ij}(x) E_j, \qquad a_{ij}(0) = 0.$$

One can find a change of the base $Y_j = \sum_{j=1}^{k} h_{ij}(x) X_j$ such that

$$Y_i(x) = E_i + \sum_{j=k+1}^{n} b_{ij}(x) E_j, \qquad i = 1, \dots, k.$$

Indeed, it is enough to "rectify" only the horizontal coordinates of the fields $X_1, \dots, X_k$, which is possible by inversion of the $k \times k$-matrix close to the identity.

Since E_i commute, the commutators $[Y_i, Y_j]$ should be "vertical" (their projections on the horizontal coordinates must vanish). On the other hand, since X_i are in the involution, these commutators should be also horizontal. This implies that Y_i commute between themselves. $\square$

7. Differential forms

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

$$\mathscr{C}(M) \times \mathscr{X}(M) \to \mathscr{X}(M), \qquad (f, X) \mapsto fX.$$

Problem 7.1. Assuming that $f > 0$, show that the orbits of the flows F^t_{fX} and F^t_X 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) $\xi \colon S \to \mathbb{R}$ 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.

7.2. **Differential 1-forms.**

Example 7.1. Let $u \in \mathscr{C}(M)$ be a smooth function. Then the correspondence

$$a \mapsto \xi(a) = \mathrm{d}u(a)$$

defines a linear functional on each space $T_a M$: its value on a vector $v \in T_a M$ 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

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

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

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

$$\mathrm{d}u : \mathscr{X}(M) \to \mathscr{C}(M), \qquad \mathrm{d}u(fX) = f \cdot \mathrm{d}u(X) \quad \forall f \in \mathscr{C}(M). \tag{3}$$

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

The set of all maps satisfying (3) is itself a module over $\mathscr{C}(M)$ 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 7.1. The $\mathscr{C}(M)$-module of maps $\mathscr{X}(M) \to \mathscr{C}(M)$, the dual module to $\mathscr{X}(M)$, is called the module of differential 1-forms.

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

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

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

$$\sum_k f_k \, \mathrm{d}u_k \in \Omega^1(M)$$

is a differential 1-form.

Proposition 7.1. *In any coordinate chart $x\colon U \to \mathbb{R}^n$ each 1-form can be represented as $\sum f_k(x)\,\mathrm{d}x_k$ with the appropriate smooth functions f_k.*

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

$$F^*\big(\sum f_k \, \mathrm{d}u_k\big) = \sum F^* f_k \cdot \mathrm{d}F^* u_k. \tag{4}$$

This definition is self-consistent.

Problem 7.2. If $F\colon M \to N$ is a diffeo and ξ, X live on N, then

$$\langle F^*\xi, F^*X\rangle = F^*\langle \xi, X\rangle, \qquad F^*(f\xi) = F^*f \cdot F^*\xi. \tag{5}$$

Prove it.

Definition 7.2. 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 $v \in T_aM$ the value $\langle F^*\xi(a), v\rangle$ is equal to $\langle \xi(F(a)), \mathrm{d}F(a)\cdot v\rangle$, where $\mathrm{d}F(a)\colon T_aM \to T_{F(a)}N$.

This definition makes sense even if F is not a diffeomorphism. The property (4) becomes then an easy lemma.

Remark 7.1. If $A\colon P \to Q$ is a linear map between two abstract linear spaces P, Q, then there is a naturally defined dual (conjugate) map $A^*\colon Q^* \to P^*$ between the respective dual spaces, which acts in the "opposite" direction by the formula

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

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

Example 7.2. 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.

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

$$\int_a^b f(x)\,\mathrm{d}x$$

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

If $\gamma\colon [0,1] \to M$ 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 $\gamma^*\xi = f(t)\,\mathrm{d}t$, where t is the "canonical" chart on $\mathbb{R}^1 = \mathbb{R}$.

Definition 7.3. Let $\gamma\colon [0,1] \to M$, $t \mapsto x(t)$, 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 7.2. *The integral is independent on the parametrization: if $t = t(s)\colon [0,1] \to [0,1]$ is a smooth monotonously growing self-map of $[0,1]$ into itself and $\widetilde{\gamma}\colon s \mapsto \gamma(t(s))$ is a reparametrization of Γ, then*

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

A slightly more focussed inspection allows to greatly extend this result for *any* smooth map $\phi\colon [0,1] \to [0,1]$ 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 $\gamma\colon [0,1] \to M$ with the marked "start point" $\gamma(0)$ and the "end point" $\gamma(1)$.

Example 7.3. Let $\xi = \mathrm{d}f$ for $f \in \mathscr{C}(M)$. Then for any oriented oriented curve Γ the integral of ξ can be instantly computed:

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

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 $\int_\Gamma \mathrm{d}f = 0$.

Remark 7.2. 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.

7.5. **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 $L_X\colon \Omega^1(M) \to \Omega^1(M)$.

Lemma 7.1.

$$\langle L_X\xi, Y\rangle = L_X \langle \xi, Y\rangle - \langle \xi, L_X Y\rangle = X \langle \xi, Y\rangle - \langle \xi, [X,Y]\rangle,$$
$$L_X(f\xi) = f L_X\xi + (L_X f)\cdot \xi.$$

Proof. This follows from the identity (5) applied to $F = F_X^t$ as $t \to 0$. ☐

Example 7.4. What about $L_{fX}\xi$?

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

$$L_{fX} = fL_X + \langle \mathrm{d}f, \cdot\rangle X,$$

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

Using this observation, we compute

$$\begin{aligned}\langle L_{fX}\xi, Y\rangle &= L_{fX}\langle \xi, Y\rangle - \langle \xi, L_{fX}Y\rangle \\ &= fL_X\langle \xi, Y\rangle - f\langle \xi, L_XY\rangle - \langle \mathrm{d}f, Y\rangle\langle \xi, X\rangle . \end{aligned} \tag{6}$$

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.

8. 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

$$T \times \cdots \times T \to \mathbb{R}, \qquad (v_1, \dots, v_k) \longmapsto \omega(v_1, \dots, v_k),$$

$\mathbb{R}$-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*

$$\xi_1 \otimes \cdots \otimes \xi_k \colon (v_1, \dots, v_k) \longmapsto \xi_1(v_1)\cdots\xi_k(v_k)$$

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.

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

Example 8.1. Let $\delta\colon (\mathbb{R}^n)^n \to \mathbb{R}$ 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 $b\colon T^2 \to \mathbb{R}$ is any bilinear map, then the map defined by the formula $b^\ddagger(v, w) = \frac{1}{2}\big(b(v, w) + b(w, v)\big)$ is symmetric. Conversely, if $q\colon T \to \mathbb{R}$ is a quadratic function, then its polarization

$b(v,w) = \frac{1}{2}\big(q(v+w) - q(v) - q(w)\big)$ is a symmetric bilinear form. In general, the symmetrization of a k-form $b\colon T^k \to \mathbb{R}$ is obtained by the averaging over all permutations ($\mathfrak{S}_k$ stands for the group of all permutations on k symbols) of its arguments:

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

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

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

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 8.2. 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 $0 \leqslant k \leqslant n = \dim T = \dim T^*$ into an *graded exterior algebra*. We list the properties of this algebra: they can all be established by tedious but simple computation.

Definition 8.1. 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,

$$(\alpha\wedge\beta)(v) = \sum_{\sigma\in\mathfrak{S}_{k+l}} (-1)^{|\sigma|}\,\alpha\big(\sigma^{<}(v)\big)\cdot\beta\big(\sigma^{>}(v)\big) \tag{7}$$

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 8.1. In some sources the definition of the wedge product (7) 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 8.1. *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*

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

(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*

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

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"* $\xi_1 \wedge \cdots \wedge \xi_n$, *called the* volume form Υ.
(7) *All spaces* $\bigwedge^k T^*$ *for* $k > n$ *are trivial (zero-dimensional).*
(8) *If* $A \colon T \to T$ *is a linear map, then it naturally defines the "adjoint" (dual) linear transformation* $A^* \colon \bigwedge^k T^* \to \bigwedge^k T^*$ *on* k*-forms by the formula* $(A^*\omega)(v_1, \dots, v_k) = \omega(Av_1, \dots, A_k)$. *In particular,* $A^*\Upsilon = (\det A) \cdot \Upsilon$.

8.2. **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 $\mathscr{C}(M)$-module $\mathscr{X}(M) \times \cdots \times \mathscr{X}(M)$ to $\mathscr{C}(M)$:

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

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 $\mathrm{d}x_i \in T_a^*(U)$ are 1-forms linear independent in each cotangent space. It follows then that any k-form can be written in these coordinates as

$$\omega = \sum_{i_1<\cdots<i_k} a_{i_1,\dots,i_k}(x)\,\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k} \tag{8}$$

with $\binom{n}{k}$ smooth coefficients $a_{\dots}(x)\in\mathscr{C}(U)$.

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

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

$$\langle F^*\omega, X_1,\dots,X_k\rangle(a) = \langle \omega, \mathrm{d}F(a)X_1,\dots,\mathrm{d}F_k(a)X_k\rangle(F(a)).$$

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

$$F^*\omega = F^*g\cdot\mathrm{d}(F^*f_1)\wedge\cdots\wedge\mathrm{d}(F^*f_k)$$

(in the right hand side F^* means the pullback $F^*\colon\mathscr{C}(N)\to\mathscr{C}(M)$ 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,

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

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

8.3. **Recap.** Differential k-forms are antisymmetric maps of the $\mathscr{C}(M)$-module $\mathscr{X}(M)\times\cdots\times\mathscr{X}(M)$ into $\mathscr{C}(M)$. The $\mathscr{C}(M)$-module $\Omega^k(M)$ is generated by monomials $f\,\mathrm{d}g_1\wedge\cdots\wedge\mathrm{d}g_k$, $f,g_i\in\mathscr{C}(M)$, where $\mathrm{d}\colon\mathscr{C}(M)\simeq\Omega^0(M)\to\Omega^1(M)$ 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:

$$(\mathrm{d}g_1\wedge\cdots\wedge\mathrm{d}g_k)\wedge(\mathrm{d}f_1\wedge\cdots\wedge\mathrm{d}f_l)=\mathrm{d}g_1\wedge\cdots\wedge\mathrm{d}f_l.$$

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 $F\colon N\to M$, and any vector field $X\in\mathscr{X}(M)$ induces the Lie derivation(s) $L_X\colon\Omega^k(M)\to\Omega^k(M)$ 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

$$L_Xf=\langle\mathrm{d}f,X\rangle=Xf\qquad\forall f\in\mathscr{C}(M)\simeq\Omega^0(M).$$

9. **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 $\mathscr{C}(M)$). The Lie derivative $L_X\colon\Omega^k\to\Omega^k$ 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

$$L\omega\colon(X,Y_1,\dots,Y_k)\longmapsto(L_X\omega)(Y_1,\dots,Y_k).$$

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 $f\in\mathscr{C}(M)$). This in turn follows from the fact that the Lie derivative $X,Y\mapsto L_XY=[X,Y]$ is “not linear” in the first argument: $[fX,Y]=f[X,Y]-(L_Yf)X\neq f[X,Y]$. The only exception occurs for $k=0$:

$$L_{fX}g=f(L_Xg)=\langle\mathrm{d}g,fX\rangle=f\,\langle\mathrm{d}g,X\rangle\,.$$

[1]Not a tensor in the classical language.

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.

9.1. **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

$$\begin{aligned}\langle L_{fX}\omega, Y\rangle &= fL_X \langle \omega, Y\rangle - \langle \omega, [fX, Y]\rangle \\ &= fL_X \langle \omega, Y\rangle - f\langle \omega, [X, Y]\rangle + \langle \omega, X\rangle \langle \mathrm{d}f, Y\rangle . \end{aligned} \tag{9}$$

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

$$\mathrm{d}\langle \omega, fX\rangle = \mathrm{d}(f\langle \omega, X\rangle) = f\,\mathrm{d}\langle \omega, X\rangle + \langle \omega, X\rangle\,\mathrm{d}f.$$

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

$$\begin{gathered}\eta_X = L_X\omega - \mathrm{d}\langle \omega, X\rangle \in \Omega^1(M), \\ \eta(X, Y) = \eta_X(Y) = \langle L_X\omega, Y\rangle - \langle \mathrm{d}\langle \omega, X\rangle, Y\rangle . \end{gathered} \tag{10}$$

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

A simple computation shows that

$$\eta(X, Y) = L_X \langle \omega, Y\rangle - L_Y \langle \omega, X\rangle - \langle \omega, [X, Y]\rangle , \tag{11}$$

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

Proposition 9.1. *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 $\mathrm{d}f$ contains all Lie (directional) derivatives $L_X f$, and the result $\eta(a)\colon T_a^*M \times T_a^*M \to \mathbb{R}$ at each point $a \in M$ depends only on the direction (vector) $X(a) \in T_aM$.

Definition 9.1. The 2-form η will be called the *differential* of the form ω. We will provisionally denote the corresponding operator by

$$\mathrm{d}'\colon \Omega^1(M) \to \Omega^2(M)$$

to avoid confusion with $\mathrm{d}\colon \Omega^0(M) \to \Omega^1(M)$.

Theorem 9.1. *The operator* d' *possesses the following algebraic properties*:

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

Proof. The first claim is obvious. The Leibniz rule follows from the (11) and the Leibniz rule for L_X, L_Y: the terms which are not divisible by f are $L_X f \cdot \langle \omega, Y\rangle - L_Y f\,\langle \omega, X\rangle = \langle df, X\rangle\,\langle \omega, Y\rangle - \langle \mathrm{d}f, Y\rangle\,\langle \omega, X\rangle = (\mathrm{d}f \wedge \omega)(X, Y)$.

The last assertion, somewhat unexpected, follows from the representation (11).

$$\mathrm{d}'\mathrm{d}f = L_X L_Y f - L_Y L_X f - L_{[X,Y]} f = 0$$

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

9.2. **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 $X \in \mathscr{X}$, then

$$i_X \colon \Omega^1 \to \Omega^0, \qquad i_X\omega = \langle \omega, X\rangle\,,$$

and similarly,

$$i'_X \colon \Omega^2 \to \Omega^1, \qquad (i'_X)\eta = \eta(X, \bullet) = \langle \eta, X, \bullet\rangle\,.$$

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

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

The representation (10) can be rewritten now in terms of the *four* operators,

$$L_X = i'_X \mathrm{d} + \mathrm{d}' i_X \quad \text{as operator } \Omega^1 \to \Omega^2 \to \Omega^1. \tag{12}$$

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

Definition 9.2.

$$i_X \colon \Omega^k(M) \to \Omega^{k-1}(M), \qquad (i_X\omega)(Y_1, \dots, Y_{k-1}) = \omega(X, Y_1, \dots, Y_{k-1}).$$

The properties of this operator are obvious.

Theorem 9.2 (Leibniz rule for the antiderivation)**.**

$$i_X(\omega \wedge \eta) = (i_X\omega) \wedge \eta + (-1)^{\deg \omega} \omega \wedge (i_X\eta), \tag{13}$$

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 $\xi \in \Omega^1$,

$$\xi^\wedge \colon \Omega^k \to \Omega^{k+1}, \qquad \omega \longmapsto \xi \wedge \omega. \tag{14}$$

Problem 9.1. Find linear algebraic identities relating operators $\xi^\wedge$ and i_X.

10. Exterior derivative as an abstract need

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

$$\Omega^0(M) \xrightarrow{\mathrm{d}} \Omega^1(M) \xrightarrow{\mathrm{d}'} \Omega^2(M) \xrightarrow{\mathrm{d}''} \Omega^3(M) \to \cdots \to \Omega^n(M) \qquad (15)$$

satisfying the conditions of Theorem 9.1?

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

First, we can postulate the Leibniz rule and require that

$$\mathrm{d}(\omega \wedge \xi) = (\mathrm{d}\omega) \wedge \xi + (-1)^{\deg \omega} \omega \wedge (\mathrm{d}\xi) \quad \text{for any two forms } \omega, \xi. \qquad (16)$$

Theorem 10.1. *There exists a unique family of $\mathbb{R}$-linear operators*

$$\mathrm{d} = \mathrm{d}^{(k)} \colon \Omega^k(M) \to \Omega^{k+1}(M)$$

such that:

(1) $\mathrm{d}^{(0)}$ *coincides with the differential on* $\Omega^0(M)$ *and with the exterior differential* $\mathrm{d}' = \mathrm{d}^{(1)}$ *on* $\Omega^1(M)$*, in particular,* $\mathrm{d}'\mathrm{d}f = 0$ *for any* $f \in \Omega^0(M) \simeq \mathscr{C}(M)$*,*
(2) d *satisfies the Leibniz rule* (16)*.*

Proof. For any k-form

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

define

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

One can then expand each differential

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

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

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

Example 10.1. For $f \in \Omega_0$ we have $\mathrm{d}f = \sum_i \frac{\partial f}{\partial x_i}\,\mathrm{d}x_i$ and

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

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

Theorem 10.2 (The Cartan "magic" formula, a.k.a. the homotopy formula)**.** *The formula* (12) *holds for forms of any degree,*

$$L_X = \mathrm{d}i_X + i_X\mathrm{d}. \tag{17}$$

Proof. The formula is true for 1-forms, and for any monomial k-form $\omega = f\,\mathrm{d}g_1 \wedge \cdots \wedge \mathrm{d}g_k$ one can check it directly assuming that $\mathrm{d}\omega = \mathrm{d}f \wedge \mathrm{d}g_1 \wedge \cdots \wedge \mathrm{d}g_k$ and (repeatedly) the Leibniz rule for i_X and L_X. ☐

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

$$i_X\mathrm{d}\omega = L_X\omega - \mathrm{d}i_X\omega,$$

we can *define* $\mathrm{d}\omega$ 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 $\mathrm{d}\omega$ defined by this way, will be indeed a multilinear antisymmetric $(k+1)$-form.

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

$$\Omega^0(M) \xrightarrow{\mathrm{d}} \Omega^1(M) \xrightarrow{\mathrm{d}} \Omega^2(M) \xrightarrow{\mathrm{d}} \Omega^3(M) \xrightarrow{\mathrm{d}} \cdots \xrightarrow{\mathrm{d}} \Omega^n(M) \tag{18}$$

with the property $\mathrm{d}^2 = 0$.

One can explicitly express the value of $\mathrm{d}\omega$ 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 10.3.

$$\begin{aligned}\mathrm{d}\omega(X_0, X_1, \dots, X_k) = \sum_{i=0}^{k} (-1)^i L_{X_i}\big(\omega(\dots, \hat{X}_i, \dots)\big) + \\ \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], \dots, \hat{X}_i, \dots, \hat{X}_j, \dots),\end{aligned} \tag{19}$$

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

Remark 10.1. 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:

$$\begin{aligned} \{i_X, \mathrm{d}\} &= L_X, \\ \{i_X, i_Y\} &= 0, \\ \{\mathrm{d}, \mathrm{d}\} &= 0. \end{aligned} \tag{20}$$

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

$$\begin{aligned} \{L_X, L_Y\} &= L_{[X,Y]}, \\ \{L_X, i_Y\} &= i_{[X,Y]}, \\ \{L_X, \mathrm{d}\} &= 0. \end{aligned} \tag{21}$$

Problem 10.1. Compute the twisted commutators involving the operators $\xi^\wedge$, see (14).

11. The general Stokes theorem Executive summary

The raison d'être of differential forms is to be integrated over higher-dimensional analogs of smooth paths.

11.1. **Integration of forms.** One can choose a simple geometric shape, say, a cube or a simplex Δ in $\mathbb{R}^k$ as the parametrization source and define a smooth (parameterized) k-cell on a manifold M as a smooth map $\sigma\colon \Delta \to M$. The integral of a k-form ω is defined as the integral

$$\int_{\sigma(\Delta)} \omega = \int_\Delta \sigma^*\omega,$$

where the latter integral of a k-form $\sigma^*\omega = a(x)\,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_k$ over a simplex in $\mathbb{R}^k$ is by definition regarded as a Riemann integral of the (smooth) function $a\colon \Delta \to \mathbb{R}$. Defined this way, the integral is independent of the "parametrization" of the image $\sigma(\Delta)$ by σ: if $h\colon \Delta \to \Delta$ is an orientation-preserving self-diffeomorphism of Δ, then integrals over two cells, σ and $\sigma \circ h$, will be the same. This follows from the formula of change of variables in the Riemann integral.

11.2. **Combinatorial topology.** The above definition is linear with respect to $\omega \in \Omega^k(M)$, and we can make it "linear" with respect to the domain of integration: we can consider formal (finite) combinations of cells with integer coefficients and extend the construction of integral by linearity. The cell $-\sigma = (-1) \cdot \sigma$ is understood then as a cell $\sigma \circ h$, where $h\colon \Delta \to \Delta$ is any self-diffeomorphism of Δ which reverses the orientation. Integer combinations of cells are called *k-chains*.

Construction of chains is necessary to explain the notion of the *oriented boundary*. For instance, the boundary of a oriented 1-cell in general consists of two geometric points, and the Newton–Leibniz formula suggests that the endpoint should be considered with the coefficient $+1$, while the starting point with -1. In the same way the boundary of a k-simplex consists of its faces, $(k-1)$-dimensional simplices themselves, which should be properly oriented (we omit completely the discussion of what is an orientation per se). One obviously translates this into the notion of a boundary $\partial\sigma$ of a cell σ.

The construction of the oriented boundary is tailored to agree with the "linear" structure of chains:

$$\partial \left(\sum n_i \sigma_i\right) = \sum n_i \, \partial \sigma_i .$$

A simple combinatorial argument shows the oriented boundary possesses the property that initially looks surprising: $\partial\partial\Delta = 0$. The explanation comes from the rule defining orientation: every $k-2$-dimensional face of Δ enters the expression for $\partial\partial\Delta$ twice, as it belongs to exactly two faces of the $\partial\Delta$, and orientation assigns them the opposite orientations so that they cancel each other in the final count.

11.3. **The general Stokes theorem.**

Theorem 11.1. *For any k-chain σ on a manifold M and any smooth $k-1$-form $\omega \in \Omega(M)$*

$$\int_{\partial\sigma} \omega = \int_{\sigma} \mathrm{d}\omega .$$

Idea of the proof. The above functoriality, linearity and possibility of subdividing cells into sums of smaller cells implies that it is sufficient to prove this result for a particular case where $M = \mathbb{R}^k$, σ is a cube $[0,1]^k$ and ω is a monomial $k-1$-form, say,

$$\omega = a(x_1, \dots, x_k) \, \mathrm{d}x_2 \wedge \cdots \wedge \mathrm{d}x_k .$$

Then
$$\mathrm{d}\omega = \frac{\partial a}{\partial x_1}(x_1, x_2, \dots, x_k)\,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_k$$
and the result will follow from the Newton-Leibniz theorem applied to functions of a single variable x_1 with $x_2, \dots, x_n$ considered as parameters. $\square$

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 klassicheskoĭ 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

12. De Rham cohomology

12.1. **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, $\mathrm{d}\omega = 0$. We construct explicitly a $(k-1)$-form α such that $\omega = \mathrm{d}\alpha$.

Consider the *linear homotopy* contracting U to the origin. By definition, this is a one-parametric family of smooth maps $H_t \colon x \mapsto tx$, $x \in U$, $t \in [0,1]$. Obviously, $H_1^*\omega = \omega$, $H_0^*\omega = 0$. The derivative $\frac{\mathrm{d}}{\mathrm{d}t} H_t^*$ 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):

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

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

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

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.

12.2. **De Rham complex.** The family of infinite-dimensional spaces equipped with the operator $\mathrm{d}^2 = 0$. The cohomology of this complex is by definition $\operatorname{Ker} \mathrm{d} / \operatorname{Img} \mathrm{d}$, 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 stack.

12.3. **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 $f_i \in \mathscr{C}(U_i)$, such that $\mathrm{d}f_i = \omega$ in U_i. Such functions are defined modulo constant terms: $\widetilde{f}_i = f_i + c_i$ 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, $\mathrm{d}(f_i - f_j) = 0$. 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

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

Then replacing the functions f_i by the functions $\widetilde{f}_i = f_i + c_i$, we see that $\widetilde{f}_i = \widetilde{f}_j$ 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 $\widetilde{f}$ such that $\mathrm{d}\widetilde{f} = \omega$.

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 $I_2 = \{(i,j) : U_{ij} \neq \varnothing\} \subseteq I_1 \times I_1$ be the set of indices corresponding to nonempty pairwise intersections. Consider the linear map

$$\mathbb{R}^{I_1} \to \mathbb{R}^{I_2}, \qquad \{c_i\} \mapsto \{c_i - c_j\}.$$

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 12.1. Consider the circle which is covered by two maps...

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

$$\mathrm{d}f_i - \mathrm{d}f_j = \alpha_{ij} \qquad \text{in } U_{ij}.$$

Since $\mathrm{d}\alpha_{ij} = 0$, we can find smooth functions g_{ij} such that $\mathrm{d}f_i - \mathrm{d}f_j = \mathrm{d}g_{ij}$. 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

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

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

12.5. **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 12.1. A (real) k-cochain $\mathfrak{c}$ is a collection of real numbers $\mathfrak{c} = \{\mathfrak{c}_{i_0 i_1 \cdots i_k}\}$ associated with all non-empty intersections $U_{i_0 \cdots i_k} = U_{i_0} \cap \cdots \cap U_{i_k}$ which is antisymmetric with respect to all permutations of the indices.

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

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

(the rth index is omitted).

Proposition 12.1.

$$\delta\delta = 0.$$

13. **Riemannian manifolds**

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

Example 13.1. Prove this and construct an atlas of charts on TM.

Recall that each bilinear form $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ can be made into a quadratic form $Q\colon \mathbb{R}^n \to \mathbb{R}$ 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 $Q(v) \geqslant 0$ and $Q(v) = 0 \iff v = 0$.

13.1. **Riemannian manifolds: definitions.**

Definition 13.1. A Riemannian metric on a manifold M is a smooth function $g\colon TM \to \mathbb{R}_+$ which, when restricted on any tangent space T_aM, is a positive definite quadratic form. A Riemannian manifold is a manifold equipped with a specific Riemannian metric.

Example 13.2. 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 $\mathbb{T}^2_{a,b} = \mathbb{R}^2/a\mathbb{Z} + b\mathbb{Z}$ 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 13.2. 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 $F\colon (M,g)\to(N,h)$ is a differentiable injective[2] map such $F^*h=g$, then we say that M is isometrically embedded in N.

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

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

Example 13.3. The vector field $\operatorname{grad} f$ is dual to the 1-form $\mathrm{d}f$.

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 $\gamma\colon[0,1]\to M$ is the integral

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

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

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

where ξ_γ is the 1-form dual to the velocity vector 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 13.4. 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!

---

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

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

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

Theorem 13.1. *Any smooth manifold can be equipped with a Riemannian metric.*

Proof. Let $\{U_i\}$ be a locally finite covering of M by charts $x_i\colon U_i \to \mathbb{R}^n$ and $1 = \sum_i \psi_i$ the subordinate partition of unity, $\mathscr{C}(M) \ni \psi \geqslant 0$, $\operatorname{supp}\psi_i \subseteq U_i$. 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! □

14. Hypersurfaces in $\mathbb{R}^n$ as Riemannian manifolds: Gauss theory

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

$$\overline{\nabla}_X Y = (L_X Y_1, \dots, L_X Y_n), \qquad X, Y, \overline{\nabla}_X Y \in \mathscr{X}(\mathbb{R}^n), \qquad (22)$$

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

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

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

1. $\overline{\nabla}$ is $\mathbb{R}$-linear in both arguments X, Y,
2. $\overline{\nabla}_X$ respects the Leibniz rule: $\overline{\nabla}_X(fY) = (\overline{\nabla}_X f)Y + f(\overline{\nabla}_X Y)$. There is only one way to interpret $\overline{\nabla}_X f$ for $f \in \mathscr{C}(M)$ as $L_X f$.

(3) $\overline{\nabla}_{fX}Y = f\overline{\nabla}_X Y$, that is, the operator $\overline{\nabla}$ acts on vector fields in the same way as L_X acts on functions (such behavior is called *tensorial*).
(4) Moreover, if $Y, Z \in \mathscr{X}(\mathbb{R}^n)$ are two vector fields, then the Leibniz rule holds for the scalar product $\langle\cdot,\cdot\rangle$:

$$\overline{\nabla}_X \langle Y, Z\rangle = L_X \langle Y, Z\rangle = \left\langle \overline{\nabla}_X Y, Z\right\rangle + \left\langle Y, \overline{\nabla}_X Z\right\rangle. \tag{23}$$

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

$$\overline{\nabla}_X Y - \overline{\nabla}_Y X = [X, Y]. \tag{24}$$

Definition 14.1. A differential operator $\nabla\colon \mathscr{X}(M)\times\mathscr{X}(M) \to \mathscr{X}(M)$ 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 (24) holds. If M is a Riemannian manifold and (23) holds, we say about the Riemannian connexion.

The property (23) means that the parallel transport in $\mathbb{R}^n$ is an isometry between $T_{\gamma(0)}\mathbb{R}^n$ and $T_{\gamma(1)}\mathbb{R}^n$.

14.2. **Induced covariant derivation.** We will show how the derivation $\overline{\nabla}_X Y$ 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 $\overline{\nabla}_X Y$ for two vector fields $X, Y \in \mathscr{X}(M)$ tangent to M is not necessarily tangent to M.

The required correction is both natural and minimal.

Definition 14.2. The covariant derivative $\nabla_X Y$ of a vector field $Y \in \mathscr{X}(M)$ along $X \in \mathscr{X}(M)$ is the orthonormal projection of $\overline{\nabla}_X Y \in \mathscr{X}(\mathbb{R}^n)$ 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 $Y \in \mathscr{X}(M)$ is said to be parallel along a curve γ : $[0, 1] \to M$, 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 14.1. Let $E_1, \dots, E_n \in \mathscr{X}(M)$ 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,

$$\nabla_{E_i} E_j = \sum_{k=1}^{n} \Gamma_{ij}^k E_k, \tag{25}$$

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:

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

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

For any $X \in \mathscr{X}(M)$ the "ambient covariant derivative" $\overline{\nabla}_X N$ is tangent to M. Indeed, by the Leibniz rule

$$\left\langle \overline{\nabla}_X N, N \right\rangle = \tfrac{1}{2} L_X \langle N, N \rangle = 0, \qquad \text{hence} \quad N \perp \overline{\nabla}_X N \in \mathscr{X}(M).$$

Definition 14.3. The *Weingarten operator* is the linear operator

$$W_a \colon T_aM \to T_aM$$

which sends a vector v into $\overline{\nabla}_v N$ (since $\overline{\nabla}$ is a covariant derivative, one can choose any $X \in \mathscr{X}(M)$ such that $X(a) = v$, and compute $\overline{\nabla}_X N$). For a vector field $X \in \mathscr{X}(M)$ we denote by WX the field $a \mapsto W_a X(a) \in T_aM$.

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

Example 14.2. 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 14.1. *The Weingarten operator is self-adjoint in the induced metric on T_aM.*

Proof. Let $X, Y \in \mathscr{X}(M)$ 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 (23) and the symmetry condition (24) of $\overline{\nabla}$, we conclude that

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

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

Definition 14.4. 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_pM 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.

14.3. **The Gauss equation.** Using the Weingarten operator, one can easily express the induced covariant derivative as it was defined in Definition 14.2: to compute the orthogonal projection, one has to add to $\overline{\nabla}_X Y$ the normal vector N with a suitable coefficient,

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

since

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

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

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

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

$$\nabla_X = \overline{\nabla}_X - \langle WX, \cdot \rangle N,$$

applying the result to a third vector field $Z \in \mathscr{X}(\mathbb{R}^n)$ 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:

$$\left(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}\right) Z = \langle WY, Z \rangle WX - \langle WX, Z \rangle WY, \quad (26)$$

and another identity valid for any Z, which implies that

$$\nabla_X (WY) - \nabla_Y (WX) = W[X, Y]. \quad (27)$$

The identity (26) is very remarkable: it asserts that a certain differential operator that could apriori 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

$$R(X, Y, Z, V) = \left\langle \left(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}\right) Z, V \right\rangle$$
$$= \left\langle \langle WY, Z \rangle WX - \langle WX, Z \rangle WY, V \right\rangle$$
$$= \langle WY, Z \rangle \langle WX, V \rangle - \langle WX, Z \rangle \langle WY, V \rangle . \quad (28)$$

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 14.1. 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

$$R(X,Y,X,Y) = \langle WX, Y\rangle \langle WY, X\rangle - \langle WY, X\rangle \langle WY, Y\rangle \\ = -\det W = -\lambda_1\lambda_2 = -R(X,Y,Y,X) \quad (29)$$

is the Gaussian curvature of the surface.

This formula is important since the left hand side of (26) 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 $\mathbb{R}^3$.

14.5. **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 14.1. *The covariant derivation ∇ on a Riemannian manifold, which is symmetric and preserves the Riemannian structure, i.e., $\forall X, Y, Z \in \mathscr{X}(M)$ satisfying the conditions*

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

is unique.

Remark 14.1. 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

$$\langle X(1), Y(1)\rangle - \langle X(0), Y(0)\rangle = \int_0^1 \nabla_{\dot\gamma} \langle X(t), Y(t)\rangle \,\mathrm{d}t \\ = \int_0^1 \langle \nabla_{\dot\gamma} X, Y\rangle + \langle X, \nabla_{\dot\gamma} Y\rangle \,\mathrm{d}t = 0.$$

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

$$\begin{aligned} \nabla_X \langle Y, Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, \\ \nabla_Y \langle X, Z\rangle &= \langle \nabla_Y X, Z\rangle + \langle X, \nabla_Y Z\rangle, \\ \nabla_Z \langle X, Y\rangle &= \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle \end{aligned} \quad (30)$$

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

$$\begin{aligned}\nabla_X \langle Y, Z\rangle + \nabla_Y \langle X, Z\rangle - \nabla_Z \langle X, Y\rangle = \\ \langle \nabla_X Y, Z\rangle + \langle Y, [X, Z]\rangle + \langle \nabla_Y X, Z\rangle + \langle X, [Y, Z]\rangle \\ = 2\langle \nabla_X Y, Z\rangle - \langle Z, [X, Y]\rangle = 2\langle \nabla_X Y, Z\rangle . \end{aligned} \tag{31}$$

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 $E_1, \dots, E_n \in \mathscr{X}(M)$ be coordinate vector fields for any local coordinate system (commuting by definition), and apply the identity (31) 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). □

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

15. **Geodesics**

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

15.1. **Some obvious computations.** For a vector field
$$W = (w_1(x), \dots, w_n(x)) = \sum_i w_i(x)\partial_i$$
to be parallel (or *constant*) along a smooth curve $\gamma\colon t \mapsto x(t)$ with the velocity vector $\dot\gamma(t) = (\dot x_1(t), \dots, \dot x_n(t)) = (v_1(t), \dots, v_n(t))$, $t \in [0, T]$, one has to meet the condition $\nabla_{\dot\gamma} W = 0$, i.e.,
$$0 = \nabla_{\dot\gamma}\bigl[\sum_i w_i\partial_i\bigr] = \sum_i \left[(\nabla_{\dot\gamma} w_i)\partial_i + w_i \sum_{k,j} \Gamma_{ij}^k v_j \partial_k\right],$$
This vector equation yields n scalar ODEs
$$\dot w_k(t) + \sum_{ij} \Gamma_{ij}^k(t) w_i(t) v_j(t) = 0, \qquad k = 1, \dots, n.$$
This can be considered as a system of *linear* equations for the unknown functions $w(t) = (w_1(t), \dots, w_n(t))$ on $[0, T]$. For any initial condition $W(0)$ it uniquely defines the solution W along any interval. The vector $W(T)$ is the result of the parallel transport along γ from $\gamma(0)$ to $\gamma(T)$.

Definition 15.1. A curve is called *geodesic*, if its velocity is constant (along the curve).

This is the nearest approximation for the mechanical idea of a straight line as a trajectory of a particle moving in absence of any external forces.

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
$$\ddot x_k(t) = \sum_{i,j} \Gamma_{ij}^k\bigl(x(t)\bigr)\,\dot x_i(t)\dot x_j(t).$$

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

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

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 $\gamma(ct)\colon [0, c^{-1}T]$ 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

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

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

$$\exp\colon TM \times \mathbb{R} \to M, \qquad (a,v,t) \longmapsto \gamma_{a,v}(t),$$

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 $\gamma_{a,v}\colon t \mapsto M$ 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 15.1. *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 15.2 (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.*

15.2. **Exponential map and local minimality.** Consider a point $a \in M$ and the exponential map $\exp_a : T_aM \to (M, a)$. Let $S\colon (M, a) \to \mathbb{R}_+$ 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 $\mathbb{R}_+ v = \{rv\colon v \in T_aM,\ r \in \mathbb{R}_+\} \subset T_aM$ are orthogonal to the spheres.

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

Proposition 15.2. *The geodesics* $\gamma_{a,rv} = \exp_a(rv)$ *are orthogonal to the level hypersurfaces* $S = \text{const}$*.*

Proof. Consider the vector field $\mathcal{R}$ 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 $\mathcal{R}$, $[\mathcal{R}, X] = 0$. We prove that $\langle \mathcal{R}, X\rangle \equiv 0$. For that sake differentiate the identity $\langle \mathcal{R}, \mathcal{R}\rangle \equiv 1$ along X:

$$0 = \nabla_X \langle \mathcal{R}, \mathcal{R}\rangle = \langle \mathcal{R}, \nabla_X \mathcal{R}\rangle = \langle \mathcal{R}, \nabla_{\mathcal{R}} X\rangle ,$$

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

$$\nabla_{\mathcal{R}} \langle \mathcal{R}, X\rangle = \langle \nabla_{\mathcal{R}} \mathcal{R}, X\rangle + \langle \mathcal{R}, \nabla_{\mathcal{R}} X\rangle = 0 + 0 = 0,$$

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

Corollary 15.1. *For as long as the exponential map* $\exp_a : T_aM \to (M, a)$ *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 $\mathrm{d}r$ 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. □

15.3. **Variations theory for geodesics.** Given a smooth curve $\gamma\colon [0,T]\to M$ on a Riemannian manifold, one can define two closely related integrals, *length* and *action*,

$$L(\gamma)=\int_0^T \|\dot\gamma(t)\|\,\mathrm dt,\qquad E(\gamma)=\int_0^T \|\dot\gamma(t)\|^2\,\mathrm dt,$$

where $\|v\|^2=\langle v,v\rangle$ is the Riemannian scalar square. The first integral has an advantage that it does not depend on the parametrization of the curve γ, only on its image. However, the integrand in it is a non-smooth function, unlike in the second case. By the Cauchy inequality,

$$L^2(\gamma)=\biggl(\int_0^T \|\dot\gamma\|\cdot 1\,\mathrm dt\biggr)^2\leqslant \int \|\dot\gamma\|^2\,\mathrm dt\cdot\int_0^T \mathrm dt=E(\gamma)T.$$

Consider a *one-parametric family of smooth curves*

$$[0,1]\times(\mathbb R^1,0)\ni(t,s)\mapsto\boldsymbol\gamma(t,s)=\gamma_s(t)\in M,$$

which we consider as a deformation of the smooth curve γ_0. Consider the function of s defined as $E(\gamma_s)\colon(\mathbb R^1,0)\to\mathbb R$. It defines two *commuting* vector fields which are $\boldsymbol\gamma_*$-images of the fields ∂_t and ∂_s on $[0,T]\times(\mathbb R^1,0)$ (for simplicity we assume that $\boldsymbol\gamma$ is injective almost everywhere).

Denote these fields by V (the "velocity" of trajectories) and W (the "variation field). Together with V we consider also the "acceleration" field $A=\nabla_V V$.

Proposition 15.3.

$$\frac{\mathrm d}{\mathrm ds}\bigg|_{s=0}E(\gamma_s)=2\,\langle V,W\rangle\,\bigg|_{\gamma_0(0)}^{\gamma_0(1)}-2\int_0^1\langle W,A\rangle\ \mathrm dt.$$

Proof. This is nothing but the formula of derivation under the sign of integral:

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} E(\gamma_s) = \int_0^1 \nabla_W \langle V, V\rangle \,\mathrm{d}t = 2\int_0^1 \langle \nabla_W V, V\rangle \,\mathrm{d}t$$

$$= 2\int_0^1 \langle \nabla_V W, V\rangle \,\mathrm{d}t \qquad \text{because of the symmetry}$$

$$= 2\int_0^1 \nabla_V \langle V, W\rangle - \langle W, \nabla_V V\rangle \,\mathrm{d}t$$

$$= 2\langle V, W\rangle \bigg|_{\gamma_0(0)}^{\gamma_0(1)} - 2\int_0^1 \langle W, A\rangle \,\mathrm{d}t,$$

since $\nabla_V \langle V, W\rangle$ is a "full derivative" in t. □

If the variation keeps the endpoints, then the field W vanishes at these points and the first term above disappears.

Theorem 15.3. *The functional E achieves a critical point on γ_0, if and only if $A = \nabla_V V$ vanishes identically, i.e., γ_0 is a geodesic.*

Proof. If not, then one can always construct a perturbation such that $W = \rho(t)A$ with $\rho(0) = \rho(1) = 0$ and $\rho(t) \geqslant 0$, which would yield a nonzero value to the integral. □

15.4. **Second variation.** In a similar way one can consider two-parametric perturbation producing two fields W_1, W_2 along the curves, and consider the corresponding bilinear form.

The computation yields

$$\frac{\partial^2 E}{\partial s_1 \partial s_2}\bigg|_{s_1=s_2=0} = -2\int_0^1 \nabla_{W_2} \langle W_1, \nabla_V V\rangle \,\mathrm{d}t$$

$$= -2\int_0^1 \langle \nabla_{W_2} W_1, \nabla_V V\rangle + \langle W_1, \nabla_{W_2}\nabla_V V\rangle \,\mathrm{d}t$$

$$= -2\int_0^1 \langle W_2, \nabla_{W_1}\nabla_V V\rangle \,\mathrm{d}t \quad \text{since on the geodesic curve } \nabla_V V = 0.$$

But we have the identity (symmetry of the connection)

$$\nabla_{W_1} V = \nabla_V W_1,$$

and by definition of the Riemann curvature,

$$\nabla_{W_1}\nabla_V V - \nabla_V \nabla_{W_1} V = R(V, W_1)V$$

(recall that all fields commute). Therefore

$$\nabla_{W_1}\nabla_V V = R(V, W_1)W + \nabla_V \nabla_{W_1} V = R(V, W_1)W + \nabla_V \nabla_V W_1$$

$$\frac{1}{2}\frac{\partial^2 E}{\partial s_1 \partial s_2}\bigg|_{s_1=s_2=0} = -\int_0^1 \big\langle W_2, \nabla_V^2 W_1 + R(V, W_1)V \big\rangle \, \mathrm{d}t.$$

Remark 15.1. The integral

$$\int_0^1 \big\langle W_2, \nabla_V^2 W_1 + R(V, W_1)V \big\rangle \, \mathrm{d}t$$

is a bilinear form on vector fields W_1, W_2 along the geodesic curve γ. It is in fact *symmetric* (as the Hessian of a smooth functional $E(\cdot)$, but this is completely non-obvious from its explicit expression.

15.5. **Degeneracy of the action.** One can easily get convinced that on small enough curves the action functional is positive definite: its Hessian (computed above) is positive definite quadratic form on vector fields along any sufficiently short geodesic curve. The situation may change if we consider longer curves.

Definition 15.2. A vector field J along a geodesic curve γ is called the Jacobi field, if

$$\nabla_V^2 + R(V, J)V = 0.$$

In coordinates it becomes a second order linear (matrix) equation of the form

$$\ddot{J}(t) = A(t)J(t), \quad A(t) = R(\,\cdot\, V(t))V(t), \; J(t) \in \mathbb{R}^n, \; A(t) \in \operatorname{Mat}_n(\mathbb{R}),$$

and it has $2n$ linear independent solutions on any interval.

A Jacobi field with zero boundary conditions $J(0) = J(1) = 0$, if it exists, implies that the (infinite dimensional) quadratic form E'' is degenerate (and this raises suspicion that it might not yield the minimum to the action). In fact, the Jacobi field with the initial condition $J(0) = 0$ can be seen as the linearization of the equation $\nabla_V V = 0$ for geodesics.

15.6. **Pencil of geodesics.** Assume that we have a one-parametric family of curves γ_s, all defined on $[0, 1]$ and starting at a common point $a = \gamma_s(0)$ for all $s \in (\mathbb{R}^1, 0)$. Then the vector field

$$W = \frac{\partial \gamma_s}{\partial s}\bigg|_{s=0}$$

along γ_0 is a Jacobi field. Indeed, since $\nabla_V V \equiv 0$ for all t, s, we can differentiate it in s, and again using symmetry of ∇ and the definition of R, conclude that

$$0 = \nabla_W \nabla_V V = \nabla_V \nabla_W V + (RV, W)V = \nabla_V \nabla_V W + R(V, W)V.$$

In fact, this computation does not depend on whether the deformation fixes the endpoints or not.

Thus the curvature tensor is responsible for the behavior of geodesics infinitely close to V: if it is negative, then they spread exponentially from each other, otherwise there are focal points where geodesics get focused after initially going in different directions.

16. **Lie groups and Lie algebras**

16.1. **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 $G \times G \to G$, $(g, h) \mapsto gh^{-1}$ is smooth. One can instantly see that in this case the maps $l_g \colon h \mapsto gh$, $r_g \colon h \mapsto hg$ 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!).

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

We list the principal examples of (finite-dimensional) Lie groups[4].

(1) Linear spaces $\mathbb{R}^n$ with the usual structure of Abelian additive group.
(2) The multiplicative groups $\mathbb{C}^* = \mathbb{C} \smallsetminus \{0\}$ 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* $\mathrm{GL}(n, \mathbb{R})$ of real nondegenerate $n \times n$-matrices.
(5) The *special linear group* $\mathrm{SL}(n, \mathbb{R}) = \{M \in \mathrm{GL}(n, \mathbb{R}) : \det M = 1\}$ of matrices with the unit determinant.
(6) The *orthogonal group* $\mathrm{O}(n, \mathbb{R}) = \{M \in \mathrm{GL}(n, \mathbb{R}) : \langle Mx, My \rangle = \langle x, y \rangle\}$ of orthogonal matrices preserving the Euclidean structure on $\mathbb{R}^n$ (rigid rotations around the origin and reflections).
(7) The s*pecial orthogonal group* $\mathrm{SO}(n, \mathbb{R}) = \{M \in \mathrm{SO}(n, \mathbb{R}) : \det M = 1\}$ 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 $\mathbb{R}^n$.

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

[4]Infinite-dimensional Lie groups are also immensely important.

(1) Cartesian (direct) product.
(2) Semidirect product (amalgam). E.g., the group $\mathrm{SO}(n,\mathbb{R})\ltimes\mathbb{R}^n$ 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.

16.4. **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 $l_g\colon h\mapsto gh$. 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 $\mathrm{d}l_{hg^{-1}}\colon T_gG\to T_hG$ 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 $v\in T_eG$ tangent to G at the unit point of the group can be carried out to any point g by the differential $\mathrm{d}l_g(v)\in T_gG$, together forming the vector field X which is *left invariant*, $\mathrm{d}l_gX=X$ 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 $X(e)\in T_eG$. Thus the space of all invariant vector fields is finite-dimensional and isomorphic to T_eG. The operator of commutator becomes then a bilinear non-commutative operator on T_eG, satisfying the standard conditions

$$[X,Y]=-[Y,X],\qquad [X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]. \tag{32}$$

Definition 16.1. A finite-dimensional linear space $\mathfrak{g}$ equipped with a bilinear operation (32), 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.

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

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

Let $X_1,\dots,X_n \in \mathfrak{g}$, $n = \dim G = \dim g$ 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.

16.6. **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.

16.7. **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

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

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

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

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, $\omega_i(X_j) = \delta_{ij}$.

16.8. **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 16.1. 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 16.1. *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 $\phi\colon \mathfrak{g}\mathfrak{h}$ extends to a Lie group homomorphism (smooth map preserving the group structure) $f\colon G \to H$ such that $\mathrm{d}f = \phi$.*

16.9. **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 $\frac{\mathrm{d}}{\mathrm{d}t}$ into $X(e)$, so that we have $\dot{g}(t) = \mathrm{d}l_{g(t)}X(e)$.

Definition 16.2. The exponential map $\exp\colon \mathfrak{g} \to G$ is the map that sends each vector $X \in T_eG \simeq \mathfrak{g}$ into the point $\exp X = g(1)$, where $g(\cdot) : \mathbb{R} \to G$ is the above one-parameter subgroup.

Theorem 16.2. *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: $\exp\colon \mathbb{R}^1 \to \mathbb{T}^1$, $\exp x = \mathrm{e}^{2\pi\mathrm{i}x}$.

For the matrix group $\mathrm{GL}(n, \mathbb{R})$ 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 $\mathrm{e}^{tx}\mathrm{e}^{sx} = \mathrm{e}^{(t+s)x}$. The last identity implies that $\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \exp tX = X$, that is, the exponential series indeed is tangent to the prescribed velocity vector X.

16.10. **About notations.** The readers have a full right to be frustrated by the notation: indeed, the square braces $[X, Y]$ in these notes were used for several seemingly unrelated things:

(1) Commutator of two elements in a an arbitrary group or algebra, $[X, Y] = XY - YX$, say, in a matrix group;
(2) Commutator of two vector fields $X, Y \in \mathscr{X} = \mathscr{X}(M)$;
(3) The bracket operation in a Lie algebra $\mathfrak{g}$.

However, these things are actually the same.

Consider the *infinite-dimensional* Lie group $\mathrm{Diff}(M)$ of C^∞-smooth self-diffeomorphisms of a manifold M with the operation of composition. Then the tangent space to this group at the identity is the space $\mathscr{X}(M)$ of smooth vector fields: if $\gamma\colon (\mathbb{R}^1, 0) \to \mathrm{Diff}(M)$, $t \mapsto F_t$, $F_0 = \mathrm{id}$ is a smooth curve in the group through the origin, then the corresponding vector field at a point $a \in M$ is the vector $X(a) \in T_aM$ tangent to the curve $t \mapsto F_t(a)$ on M.

The exponential map will then be flow map,

$$\exp\colon X \mapsto \exp X = F^1, \qquad \text{where } \{F_X^t\}_{t\in\mathbb{R}} \text{ is the flow map of } X.$$

It is not surjective anymore, since the "manifold" $\mathrm{Diff}(M)$ is infinite-dimensional. Thus $\mathscr{X}(M)$ is an infinite-dimensional Lie algebra, with the bracket defined by the commutator of the vector fields.

In a similar way we can consider the infinite-dimensional vector space $\mathscr{C}(M)$ and the infinite-dimensional Lie subgroup $G = \mathrm{Aut}\,\mathscr{C}(M)$ of $\mathbb{C}$-linear automorphisms of this algebra inside the "general linear" group $\mathrm{GL}(\mathscr{C}(M))$ of linear invertible self-maps of the $\mathbb{C}$-space $\mathscr{C}(M)$. The Lie algebra $\mathfrak{g}$ of G consists of derivations of $\mathscr{C}(M)$, and exactly as in subgroups of $\mathrm{GL}(n, \mathbb{R})$ the bracket is defined by the commutator of derivations $[X, Y] = XY - YX$.

However, a systematic treatment of infinite-dimensional Lie groups is far beyond the scope of these notes.

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

17. **Symplectic manifolds and Hamiltonian dynamics**

17.1. **First vs. second order equations.** The Newtonian laws of mechanics have the form of ordinary differential equations, but of second (rather than the first) order: the forward behavior of a system is determined by its initial location and the initial velocity, very much like the equation for geodesics (which are in fact equations of a free motion in curved space without acting forces). In presence of forces the equations become more complicated but retain the second order.

This means that the natural phase space for mechanical equation should be the tangent bundle TN of a smooth manifold (the *configuration space*), which is even-dimensional. The corresponding equations should take the form then

$$\dot x_i = v_i, \qquad \dot v_i = F_i(x,v), \qquad i=1,\dots,n$$

(in the non-autonomous case F_i may explicitly depend on time). However, not every vector-function $F=(F_1,\dots,F_n)$ may appear in the right hand side: the Newton laws restrict very strongly the admissible things (think about conservation of energy, for instance, if the system has no friction).

To avoid possible confusion: here and below we consider only *conservative* mechanical system, without friction, in which the full energy (what it is?) is preserved.

It turns out that rather than studying vector fields on the tangent bundle, it is much more natural to consider a dual space, vector fields on the cotangent bundle T^*N which possesses an amazing extra structure, called *symplectic form*.

17.2. **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 $\pi\colon T^*N \to N$ on the base. In the local coordinates $x_1,\dots,x_n$ the coordinates on T^*_xN 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 $\pi\colon M\to N$ be be above projection. Then for any point $a=(x,\xi)\in M$, $\xi\in T_xN$, the differential $\mathrm d\pi : T_aM \to T_aN$ 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:

$$\alpha = \sum_{i=1}^{n} p_i \, \mathrm{d}x_i.$$

Definition 17.1. The *canonical symplectic structure* on T^*M is the differential 2-form $\omega = \mathrm{d}\alpha$. In the local canonical coordinates,

$$\omega = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}x_i. \tag{33}$$

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

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

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

Definition 17.2 (main)**.** A symplectic manifold is a manifold equipped with a nondegenerate 2-form ω, called symplectic structure.

Remark 17.1. Since $\det(-A) = (-1)^n \det A$, an odd-sized antisymmetric matrix cannot be nondegenerate, hence a symplectic manifold must be even-dimensional.

Example 17.1. The simplest example of a symplectic manifold is the plane with the (oriented) area as the symplectic structure.

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

Definition 17.3. A vector field $X \in \mathscr{X}(M)$ on a symplectic manifold is called *locally Hamiltonian*[5], if $\mathrm{i}_X\omega = \omega(X, \cdot) \in \Omega^1(M)$ is closed. The field is called *Hamiltonian*, if $\mathrm{i}_X\omega = \mathrm{d}H$ is exact. The function $H \in \mathscr{C}(M)$ is called the *Hamiltonian*[6] of the field X.

[5] The word *Hamiltonian* is used as an adjective here.

[6] And here *Hamiltonian* is a noun. Sorry about that.

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.

Example 17.2. If $\omega = \sum \mathrm{d}p_i \wedge \mathrm{d}x_i$, then for a Hamiltonian $H = H(x,p)$ the Hamiltonian vector field is given by the equations.

$$\dot{x} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial x}.$$

In particular, if $H(x,p) = \frac{1}{2}\langle p, p\rangle = \frac{1}{2}\sum p_i^2$ is a positive definite quadratic form then

$$\dot{x} = p, \qquad \dot{p} = 0.$$

This is the equation of a unit mass in the empty Euclidean space, if we interpret p as the vector of momenta.

More generally, on every Riemannian manifold N the Riemann structure defines a positive definite quadratic form $g = \langle v, v\rangle$ on each tangent space T_xN, which in turn induces an isomorphism between T_xN and T_x^*N and allows to bring g as a positive definite quadratic form $K(p)$ on T_x^*N. This form, considered as a Hamiltonian, gives rise to the system of the second order, whose solutions will be geodesic curves on N.

Remark 17.2. Passage from g to K has nothing to do with the differential structure on the manifold, this is purely linear algebraic transform called the Legendre transform. It associates with a positive definite quadratic form $q(v) = \frac{1}{2}\langle Av, v\rangle$ on a Euclidean space, another form $K(p) = \max_v \langle p, v\rangle - q(v)$. The maximum of the right hand side is achieved if $p = Av$, i.e., when $v = A^{-1}p$ and is equal to $\frac{1}{2}\langle A^{-1}p, p\rangle$.

Problem 17.1. Write the equation for the case where $H(x,p) = \frac{1}{2}\langle p, p\rangle + U(x)$, where U is a function of the x-variables only (the *potential*).

17.4. **Conservation laws.**

Theorem 17.1. *The flow of any locally Hamiltonian vector field preserves the symplectic structure: if* $\mathrm{d}\mathrm{i}_X\omega = 0$, *then* $L_X\omega = 0$.

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

Proof.

$$L_X\omega = \mathrm{i}_X\mathrm{d}\omega + \mathrm{d}\mathrm{i}_X\omega = 0. \quad \square$$

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

Corollary 17.1. *The flow of any locally Hamiltonian vector field is volume-preserving.* $\square$

Theorem 17.2. *The Hamiltonian function H is preserved by the corresponding Hamiltonian flow.*

Proof.

$$L_X H = \mathrm{i}_X \mathrm{d}H = \mathrm{i}_X \mathrm{i}_X \omega = \omega(X, X) = 0. \quad \square$$

17.5. **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 $X = \operatorname{sgrad} H$, $Y = \operatorname{sgrad} G$,

$$\omega(X, Y) = \mathrm{i}_X \mathrm{i}_Y \omega = L_X G = -L_Y H.$$

Definition 17.4. The above function is called the *Poisson bracket* of H, G and denoted by $\{H, G\}$.

The elementary computation yields

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

This implies that the Poisson bracket satisfies the Jacobi identity,

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

Example 17.3. 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 17.3. *The algebra of functions $\mathscr{C}(M)$ 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 $\mathscr{X}(M)$ with the commutator $[\,\cdot\,,\,\cdot\,]$ as the Lie bracket. The map $H \mapsto \operatorname{sgrad} H$ is a homomorphism of Lie algebras, whose kernel consists of constants, if M is connected.*

17.6. **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 $X = \operatorname{sgrad} H$ 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. However, the new integrals may well turn out to be functionally dependent with the previously known integrals.

Yet there is a preciously rare case of completely integrable systems, which have the maximal number of the first integrals.

Definition 17.5. A Hamiltonian vector field X with the Hamiltonian function H on a $2n$-dimensional symplectic manifold is called *completely integrable*, if there are n functions $f_1, \dots, f_n$, $f_1 = H$, with linear independent at every point differentials, such that $\{H, f_i\} = 0$.

If a completely integrable system happens to live on the simple space $\mathbb{R}^{2n}$ and a generic level surface $S_I = \{f_1 = I_1, \ \cdots, \ f_n = I_n\}$ is compact, then one can claim very strong statement about the dynamics of the field X:

- Each S_I, $I \in \mathbb{R}^n$, is a torus; the values of the functions f_i can be chosen as coordinates on M^{2n};
- One can find a local tuple of coordinates $\varphi_i \in \mathbb{R} \bmod \mathbb{Z}$, such that the collection (I, φ) is a symplectic coordinate system, i.e., $\omega = \sum \mathrm{d}I_i \wedge d\varphi_i$; the I-coordinates are called *actions*, the respective variables φ, naturally, *angles*.
- In these coordinates X is a Hamiltonian vector field with the Hamiltonian $H = H(I)$ depending only on the action variables I; the flow of X is given by the differential equations

$$\dot{I}_i = 0, \quad i =, \dots, n, \qquad \dot{\varphi}_i = \rho_i(I), \quad i = 1, \dots, n.$$

The last set of equations describes a *quasiperiodic* motion on the tori, whose dynamics depends on the existence of arithmetic relationships between the respective frequences $\rho(I)$. For a generic value of I the collection $(\rho_1(I) : \cdots : \rho_n(I))$ independent over $\mathbb{Q}$, so the trajectories are everywhere dense on the torus. *Resonances* may change this behavior.

17.7. **KAM–theory.** A small perturbation of a completely integrable Hamiltonian $H = H(I)$ in general destroys its symmetries (first integrals), although the system remains Hamiltonian. It turns out, and this was discovered in various settings by Kolmogorov, Arnold and Moser, that the majority of the invariant tori, described above, survive if the perturbation is sufficiently small, and only resonant and near-resonant tori may disappear. This has very important consequences for the stability of celestial mechanics and at the same time explains the origin of stochastic-like ("chaotic") behavior.

17.8. **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 17.4. *Any symplectic space of even dimension* $2n$ *always has a basis*

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

such that

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

Therefore all symplectic spaces of the same dimension are isomorphic.

This theorem is analogous to the theorem about existence of an orthonormal basis for a symmetric bilinear form (scalar product). However, there is an enormous difference between the symplectic and Riemannian case. However, in the symmetric case we know that it is in general impossible to find local coordinates $x_1,\dots,x_n$ on a Riemannian manifold in such a way that the metric will take the form $\sum \mathrm{d}x_i^2$: the latter is flat (zero curvature), which might not be the case for the initial metric.

In the symplectic case there is no similar obstruction.

Theorem 17.5 (G. Darboux)**.** *All symplectic structures are locally equivalent to the structure defined by the form* $\sum \mathrm{d}p_i \wedge \mathrm{d}x_i$ *as above.*

The proof of this result is a two-liner, perhaps one day I will copy it here.

This local statement implies, among other things, that two, say, topological disks on the plane can be mapped one to the other by a symplectic (area preserving) map if and only if their areas are equal.

This might lead to a (wrong) conclusion that the symplectic geometry has no invariants beyond obvious ones. In 1985 M. Gromov discovered a surprising fact, called the Non-squeezing theorem: *A ball of unit radius in the standard symplectic space* $\mathbb{R}^{2n}$ *cannot be mapped into a sufficiently thin cylinder*.

There are more things in heaven and earth, Horatio,
Than are dreamt of in your philosophy.

This result paved a way to the immensely rich theory which is called today *Symplectic Topology*.

References

[A] V. I. Arnol′d, *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

18. **Complex manifolds**

18.1. **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 $\pm\mathrm{i}$, but it impossible to distinguish between these Siamese twins: the swap $\mathrm{i}\mapsto-\mathrm{i}$ induces the (Galois) automorphism of $\mathbb{C}/\mathbb{R}$, called the complex conjugacy:

$$\overline{x+\mathrm{i}y}=x-\mathrm{i}y,\qquad x,y\in\mathbb{R}.$$

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 $z=x+\mathrm{i}y$ and $\bar z=x-\mathrm{i}y$, using the transformation $x=\frac12(z+\bar z)$, $y=\frac{\mathrm{i}}{2}(z-\bar z)$. 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

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

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

18.2. **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 $J\colon\mathbb{R}^{2n}\to\mathbb{R}^{2n}$ such that $J^2=-\operatorname{id}$. Presence of such operator allows to define the complex multiplication

$$\mathbb{C}\times\mathbb{R}^{2n}\to\mathbb{R}^{2n},\qquad (a+\mathrm{i}b)v=av+bJv,\quad v\in\mathbb{R}^{2n}.$$

A map $A\colon\mathbb{R}^{2n}\to\mathbb{R}^{2n}$ 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

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

The operator J has $2n$ eigenvalues $\pm\mathrm{i}$, 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 18.1. Let $n = 1$. Then $V = \mathbb{C} \simeq \mathbb{R}^2$ and in the coordinates (x, y) such that $z = z + \mathrm{i}y$ 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 $x = \frac{1}{2}(z+w)$, $y = \frac{\mathrm{i}}{2}(z-w)$, so that $z = x + \mathrm{i}y$ and $w = x - \mathrm{i}y$. In the new coordinates the matrix of J becomes diagonal, $J = \begin{pmatrix} \mathrm{i} & \\ & -\mathrm{i} \end{pmatrix}$.

Remark 18.1. 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 $\bar{\mathrm{i}} = -\mathrm{i}$, thus any unification of real and complex structures must be invariant with respect to the automorphism $\alpha \mapsto \bar{\alpha}$ of $\mathbb{C}$.

Remark 18.2. 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 $\pm\mathrm{i}$. This splits $V^{\mathbb{R}\mathbb{C}}$ into the direct sum of two eigenspaces, $V^{\mathbb{R}\mathbb{C}} = V' + V''$, where

$$V' = \{v \in V^{\mathbb{R}\mathbb{C}} \colon Jv = \mathrm{i}v\}, \quad V'' = \{v \in V^{\mathbb{R}\mathbb{C}} \colon Jv = -\mathrm{i}v\}.$$

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

$$\bar{\cdot}\colon V^{\mathbb{R}\mathbb{C}} \to V^{\mathbb{R}\mathbb{C}}, \qquad (z, \bar{z}) \mapsto (\bar{z}, z).$$

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

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

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

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, $\langle u, u \rangle \geqslant 0$, $\langle u, u \rangle = 0 \iff u = 0 \in \mathbb{C}^n$, 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 $\mathbb{C}^n$,

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

18.4. **Dual spaces and their decomposition.** For an $\mathbb{R}$-linear even-dimensional space $\mathbb{R}^{2n}$ the usual dual, the space if $\mathbb{R}$-linear maps $\mathbb{R}^{2n} \to \mathbb{R}$, 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 $\xi\colon V \to \mathbb{C}$,

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

(suffices to requite these conditions for $\alpha = \mathrm{i}$). 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 $z_1, \dots, z_n$, $x_i + \mathrm{i}y_i = z_i \in \mathbb{C}$, then the (constant) forms $\mathrm{d}z_i = x_i + \mathrm{i}\mathrm{d}y_i$ and $\mathrm{d}\bar{z}_i = \mathrm{d}x_i - \mathrm{i}y_i$ constitute the basis of $\mathbb{C}^{n*}$.

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

18.5. **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 $\xi_i \in V^*_{1,0}$, $\eta_j \in V^*_{0,1}$.

18.6. **Holomorphic functions and maps.** A map $f \colon U \to \mathbb{C}^m$, $U \subseteq \mathbb{C}^n$, is holomorphic at a point $a \in U$, if its differential (tangent map) $\mathrm{d}f \colon \mathbb{C}^n \to \mathbb{C}^m$ 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 $\mathrm{GL}(n, \mathbb{C})$). Composition of holomorphic maps is again holomorphic. The space of holomorphic maps $U \to \mathbb{C}$ (holomorphic functions) is denoted by $\mathscr{O}(U)$. A basic theorem of complex analysis claims that $\mathscr{O}(U) \subset \mathscr{C}(U)$, that is, existence of "just one derivative" ($\mathbb{C}$-linearity of $\mathrm{d}f$) implies infinite differentiability of the latter.

18.7. **Holomorphic manifolds and complex differential forms on them.** A *holomorphic manifold* is a topological space M equipped with an atlas of charts (homeomorphisms) $z_\alpha \colon U_\alpha \to \mathbb{C}^n$ 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, $T^*_a M = \mathrm{Hom}_{\mathbb{R}}(T_a M, \mathbb{C})$, $\dim_{\mathbb{R}} T^*_a M = 2 \dim_C T^*_a M = 4n$.

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

$$\xi = \sum_{i=1}^n a_i(z)\,\mathrm{d}z_i + b_i(z)\,\mathrm{d}\bar{z}_i, \qquad a_i, b_i \in \mathscr{C}(U)$$

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

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

Ona 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.

18.8. **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 $f(z)=u(z)+\mathrm{i}v(z)\in\mathscr{C}(U)$, then $\mathrm{d}f=\mathrm{d}u+\mathrm{i}\mathrm{d}v$, if $u,v\colon U\to\mathbb{R}$ are C^∞-smooth real and imaginary parts of a complex valued function $f\colon U\to\mathbb{C}$.

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

Definition 18.1. Consider the following differential operators on $\mathscr{C}(U)$, which in the coordinates $(z_1,\dots,z_n)$, $z_i=x_i+\mathrm{i}y_i$, take the form

$$\partial_i=\frac{\partial}{\partial z_i}=\frac12\biggl(\frac{\partial}{\partial x_i}-\mathrm{i}\frac{\partial}{\partial y_i}\biggr),\qquad \bar\partial_i=\frac{\partial}{\partial\bar z_i}=\frac12\biggl(\frac{\partial}{\partial x_i}+\mathrm{i}\frac{\partial}{\partial y_i}\biggr).$$

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

Proposition 18.1 (Cauchy–Riemann conditions)**.** *A function $f\in\mathscr{C}(U)$ is holomorphic, $f\in\mathscr{O}(U)$, if and only if $\bar\partial_i f=0$ for all $i=1,\dots,n$.*

Proposition 18.2. *The forms $\mathrm{d}z_i$ and $\mathrm{d}\bar z_i$ are dual to the vector fields $\partial_i,\bar\partial_i$:*

$$\mathrm{d}z_i(\partial_j)=\mathrm{d}\bar z_i(\bar\partial_j)=\delta_{ij},$$

with the other values vanishing, $\mathrm{d}z_i(\bar\partial_j)=\mathrm{d}\bar z_i(\partial_j)=0$ for all i,j.

Proposition 18.3. *For a smooth function $f\in\mathscr{C}(U)$ its differential $\mathrm{d}f\in\Omega^1(U)$ can be split as the sum of $\Omega^{1,0}(U)$ and $\Omega^{0,1}(U)$ parts,*

$$\mathrm{d}f=\partial f+\bar\partial f,\qquad \partial f=\sum_1^n\partial_i f\,\mathrm{d}z_i,\quad \bar\partial f=\sum_1^n\partial_i f\,\mathrm{d}\bar z_i.$$

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 $p \leqslant n$, $q \leqslant n$.

18.9. **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 $T'_aM \simeq T_aM$ 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

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

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

19. **Survey of further issues**

19.1. **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.

19.2. **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 19.1. For $n = 1$ a 1-form $\omega = a(z,\bar z)\,\mathrm{d}z + b(z,\bar z)\,\mathrm{d}\bar z$ can be integrated along curves in U. If $\omega \in \Omega^{1,0}(U)$, i.e., $b \equiv 0$ then $\mathrm{d}\omega = \bar\partial\omega = \frac{\partial}{\partial\bar z}a\,\mathrm{d}z \wedge \mathrm{d}\bar z$. If ω is holomorhic, i.e., $\bar\partial a \equiv 0$, then $\mathrm{d}\omega = 0$ 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.

19.3. **Poincaré–Grothendieck theorem.** Since $\mathrm{d} = \partial + \bar\partial$, we have $0 = \mathrm{d}^2 = \partial^2 + \partial\bar\partial + \bar\partial\partial + \bar\partial^2$. 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,

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

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

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

Theorem 19.1. *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 $q \geqslant 1$ is such that $\partial\bar\omega = 0$, then there exists $\alpha \in \Omega^{p,q-1}(U)$ such that $\omega = \bar\partial\alpha$.*

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

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

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.

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

Definition 19.1. A complex manifold is called *Hermitian*, if there is a Hermitian inner product defined on its *holomorphic* tangent bundle $TM = \bigcup_{a\in M} T_aM$, $\dim_{\mathbb{C}} T_aM = \dim M$.

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

$$\sum_{i,j} h_{ij}(z, \bar z)\,\mathrm{d}z_i \otimes_{\mathbb{C}} \mathrm{d}\bar z_j$$

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 $\omega = \frac{\mathrm{i}}{2}(h - \bar{h})$ 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

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

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

Definition 19.2. A Hermitian manifold is called a *Kähler manifold*, if the fundamental form ω is closed, $\mathrm{d}\omega = 0$ and hence defines a symplectic structure on M.

Example 19.2. 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):

$$\mathrm{d}\omega = 0$$
$$\nabla\omega = 0.$$

19.6. **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 19.3. An almost complex structure on an even-dimensional manifold is the operator field $J = \{J_a \colon T_aM \to T_aM\}$ (tensor) which apriori is variable (depending on a) satisfying the condition $J^2 = -\operatorname{id}$.

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 $\mathrm{d}\omega$ 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 $r+s=p+q+1$, $r,s\geqslant 0$, 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) $\mathrm{d}=\partial+\bar\partial$, that is, no "nonstandard" types appear by the exterior derivative,
(3) $\bar\partial^2=0$.

Theorem 19.2 (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 $J=\mathrm{i}$.*

19.7. **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 $\mathscr{O}(M)$ 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 $\mathscr{O}(M)=\mathbb{C}$.

The alternative is to work with functions holomorphic only in parts of M. Yet such an approach makes it impossible to equip $\mathscr{O}(M)$ 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

20. Exam

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

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 2. The real 2-torus $\mathbb{T}^2$ is defined as the quotient space $\mathbb{R}^2/\mathbb{Z}^2$, hence the space $\mathscr{C}(\mathbb{T}^2)$ 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 3. Let $\mathbb{D}$ be the unit disk $\{|z| < 1\} \subseteq \mathbb{C}$ with the Riemannian metric

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

1. How do you understand this definition?

2. Show that the only *invertible* holomorphic map $f : \mathbb{D} \to \mathbb{D}$ 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

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

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 $\mathbb{H} = \{\operatorname{Im} z > 0\} \subset \mathbb{C}$ be the upper half-plane and

$$\rho' = \frac{|\mathrm{d}z|}{\operatorname{Im} z}$$

a Riemannian metric on it. Show that $\mathbb{H}$ and $\mathbb{D}$ are biholomorphically equivalent, and each conformal map $f \colon \mathbb{D} \to \mathbb{H}$ 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 4. In this problem we will study behavior of geodesics on non-neatively-curved surfaces.

1. Describe all geodesics on the round sphere $\mathbb{S}^2 \subseteq \mathbb{R}^3$.

2. Let $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ be the flat torus. Describe all geodesics on $\mathbb{T}^2$. Prove that all of them are either closed (periodic), or non-selfintersecting. Does the answer change if the lattice $\mathbb{Z}$ is replaced by another lattice $a\mathbb{Z} + b\mathbb{Z}$ with $a, b \in \mathbb{R}^2$ two linear independent vectors?

3. Consider a "flat" regular tetrahedron Θ^2 (think of a paper model made of 4 equilateral triangles with some edges identified). Show that if we delete the vertices, then the remainder is a (non-complete) flat Riemannian manifold topologically equivalent to the sphere without 4 points. Construct a closed geodesic on this manifold.

4. Show that the flat tetrahedron can be 2-covered by a flat 2-torus: there exists a map $\pi \colon \mathbb{T}^2 \to \Theta^2$ such that $\pi^{-1}(a)$ consists of two points for any $a \in \Theta^2$.

Hint: consider the tiling of $\mathbb{R}^2$ by equilateral triangles and find out a parallelogram made of 8 tiles which wraps twice around Θ^2.

5. Describe all geodesics on Θ^2 which avoid vertices (to avoid ambiguities). Prove that all of them are either closed, or nonselfintersecting.

6. Consider the "flat cube" K made of 8 paper squares. Make it into a non-complete flat Riemannian manifold by deleting vertices. Compute the sum of exterior angles of any geodesic non-intersecting polygon containing one or more (deleted) vertices inside. Read somewhere about the Gauss–Bonnet theorem and explain the results.

7. Construct examples of closed geodesics on K. How many vertices they should contain inside/outside? Why?

8. Can you find a closed geodesic 6-gon on K?

9* Can you construct a closed *non-planar* geodesic 6-gon?

10*. Can you construct a self-intersecting geodesic polygon on K?

11. Construct non-closed geodesics on K.

Problem 5. Consider the space $\mathbb{R}^3$ with the "vector", or cross product, defined by the usual rules:

$$a \times b = c \iff c \perp a,\ c \perp b,\ \|c\| = \|a\|\,\|b\|\,|\sin \angle(a,b)|,$$

which leaves for c one of the two possibilities; we choose one so that the triple (a, b, c) is positively oriented.

1. Prove that this is a bilinear antisymmetric (vector-valued) operation.

2. Assume that $\mathbb{R}^3$ is equipped with the standard Euclidean structure. Find out how the exterior power $\mathbb{R}^{3*} \wedge \mathbb{R}^{3*}$ can be identified with $\mathbb{R}^{3*}$. Write the cross product through the wedge product and duality.

3. Prove by as many "independent" ways, that the cross product satisfies the Jacobi identity.

4. Prove that $\mathbb{R}^3$ with the cross product chosen for the bracket, is a Lie algebra $\mathfrak{s}$.

5. Represent this Lie algebra by a subalgebra of 3×3-matrices.

6. Find a connected Lie group S for which $\mathfrak{s}$ is a Lie algebra.

7. Explain to your closest friend-physicist how the exponential map $\exp\colon \mathfrak{s} \to \mathrm{S}$ works.

Hint: don't hesitate to use words "infinitely small", "infinitely close to" etc. Explain to him/her what means the addition in the Lie algebra $\mathfrak{s}$ is *commutative*.

Problem 6 (optional). Suggest a problem of your own for future students, which in your opinion might be instructive or fascinating.

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