Mathematical methods of classical mechanics

Lecture notes

Prologue

A good physics theory is concerned with observables, quantities that do not depend on a system of reference (that is, coordinate system and other auxiliary data, such as metric, etc). That is a lesson all mathematicians should learn too: deal only with objects that can be defined invariantly (but be always prepared to compute in coordinates).

Illustrate this philosophy with what you learned (or did you?) from Calculus: df makes invariant sense while f' does not.

Let U be some domain (manifold, if you wish), $x \in U$ a point and v a tangent vector at x. Assume a function $f : U \to \mathbf{R}$ is given. Can you make a number out of these data? Here is a construction. Let $\gamma(t)$ be a parametric curve in U s.t. $\gamma(0) = x$ and $\gamma'(0) = v$. Consider the number

$$L_v(f) := \frac{df(\gamma(t))}{dt}|_{t=0}.$$

Of course, this notation assumes that the number doesn't depend on the parameterization $\gamma(t)$. Let us see what it is in coordinates:

$$L_v(f) = \sum f_{x_i} x_i' = \nabla f \cdot v,$$

and we see that $L_v(f)$ really depends on v only. Moreover, we learn that, in coordinates,

$$L_v = \sum v_i \partial_{x_i}.$$

The operator L_v is the familiar directional derivative; we have identified tangent vectors with these operators on functions.

Now it is clear how to define a *covector* df: given a vector v, the value $df \cdot v$ is $L_v(f)$ (everything is happening at point x). And what is the gradient ∇f? It doesn't exist unless one has a Euclidean structure. If this extra structure is present then vectors and covectors are identified and ∇f is identified with df. Thus df exists while f' (gradient) does not.

1. Vector fields.

1.1. A boring but necessary exercise: how does a tangent vector change under a change of coordinates?

Let x and y be coordinate systems, v and u the *same* vector in these coordinates, respectively. The Chain Rule gives:

$$\partial_{y_j} = \sum \frac{\partial x_i}{\partial y_j} \partial_{x_i}.$$

Therefore

$$\sum v_i \partial_{x_i} = \sum u_j \partial_{y_j} = \sum u_j \frac{\partial x_i}{\partial y_j} \partial_{x_i},$$

and thus

$$v_i = \sum \frac{\partial x_i}{\partial y_j} u_j,$$

or $v = Ju$ where J is the Jacobi matrix.

Important Exercise. How does the differential df change under a change of coordinates? The answer illustrates the difference between vectors and covectors.

1.2. A vector field v is when there is a tangent vector at each point (of the domain) depending smoothly on the point. We already related a linear differential operator L_v with this field. Another object is a *1-parameter group of diffeomorphisms*, or a *flow* ϕ_t for which v is the velocity field. This means:

(i) $\phi_s \phi_t = \phi_{s+t}$;

(ii) $d\phi_t(x)/dt|_{t=0} = v(x)$.

That a vector field generates a flow is the principal theorem of the theory of ODE's.

Examples. $v = \partial_x, u = x\partial_y$ in the plane. What are the flows? More examples: $x\partial_x + y\partial_y$ (dilation), $y\partial_x - x\partial_y$ (rotation).

Exercise. Consider S^3 as the unit sphere in the space of quaternions. One has three 1-parameter groups of diffeomorphisms of the sphere: $\phi_t(x) = \exp(ti)x, \psi_t(x) = \exp(tj)x$ and $\eta_t(x) = \exp(tk)x$. Compute the respective three tangent fields in the Cartesian coordinates in the ambient 4-space.

1.3. Consider two vector fields v and u with the respective flows ϕ and ψ. Do the flows commute? Not necessarily as the above example shows. To measure the non-commutativity, consider the points $\psi_s \phi_t(x)$ and $\phi_t \psi_s(x)$. Let f be a test function. Then

$$\Delta := f(\psi_s \phi_t(x)) - f(\phi_t \psi_s(x))$$

vanishes when $t = 0$ or $s = 0$. Thus its Taylor expansion starts with the st-term.

Lemma. $\frac{\partial^2}{\partial s \partial t}(\Delta)|_{s=t=0} = (L_v L_u(f) - L_u L_v(f))(x)$.

Proof. One has:

$$\frac{df(\phi_t \psi_s(x))}{dt}|_{t=0} = (L_v f)(\psi_s(x)).$$

For $g = L_v f$ one also has:

$$\frac{dg(\psi_s(x))}{ds}|_{s=0} = (L_u g)(x),$$

thus

$$\frac{\partial^2}{\partial s \partial t}|_{s=t=0} f(\phi_t \psi_s(x)) = (L_u L_v f)(x),$$

and we are done.

The *commutator* $L_v L_u(f) - L_u L_v$ appears to be a differential operator of 2-nd order (you differentiate twice!) but, in fact, it has order 1. Here is a formula in coordinates:

$$L_v L_u - L_u L_v = \sum \Big(v_i \frac{\partial u_j}{\partial x_i} - u_i \frac{\partial v_j}{\partial x_i} \Big) \partial_{x_j}.$$

Definition. The *bracket* (or the commutator) of vector fields v and u is a new vector field $w := [v, u]$ such that $L_w = L_v L_u - L_u L_v$.

In coordinates,

$$[v, u]_j = \sum \Big(v_i \frac{\partial u_j}{\partial x_i} - u_i \frac{\partial v_j}{\partial x_i} \Big).$$

Example. Consider polynomial vector fields on the line: $e_i = x^{i+1}\partial_x,\ i \geq -1$. Then $[e_i, e_j] = (j - i)e_{i+j}$.

Exercises. 1). Continuing Exercise 1.2, compute the commutators of the vector fields from that exercise.

2). Consider a linear space and let A be a linear transformation. Then $v(x) = Ax$ is a vector field called a *linear vector field*. Show that the respective 1-parameter group consists of the linear diffeomorphisms $\phi_t(x) = \exp(tA)(x)$ where

$$e^A = E + A + \frac{A^2}{2} + \frac{A^3}{3!} + ...$$

Let B another linear map and $u(x)$ the respective linear vector field. Compute the commutator $[v, u]$. *Hint*: this is again a linear vector field.

1.4. Let v, u be vector fields and ϕ_s, ψ_t the respective flows.

Theorem. *The vector fields commute if and only if so do the flows:*

$$[v, u] = 0 \ \text{ iff } \ \psi_s \phi_t = \phi_t \psi_s.$$

Proof. In one direction the statement is clear. Outline the converse argument. Let f be a test function. We have:

$$\Delta(s, t) = f(\psi_s \phi_t(x)) - f(\phi_t \psi_s(x)) = o(s^2 + t^2),\ s, t \to 0, \tag{1}$$

and we want to show that $\Delta = 0$.

Consider a rectangle in the s, t-plane with the vertex (s_0, t_0). Divide each side into N equal parts. To a path from $(0, 0)$ to (s_0, t_0) on this lattice there corresponds the diffeomorphism according to the rule:

$$[t_1, t_2] \to \phi_{t_2 - t_1}, \quad [s_1, s_2] \to \psi_{s_2 - s_1}.$$

Any path can be changed to any other in at most N^2 elementary steps. According to (1), each such step leads to a discrepancy of order $1/N^3$. Thus Δ is of order $1/N$ for every N, that is, $\Delta = 0$.

Remark. Consider a manifold M and its submanifold N. Let u and v be vector fields on M, tangent to N. Then, obviously, u, v are vector fields on N (strictly speaking, we should denote them by $u|_N, v|_N$). It follows from our (conceptual) definition of the commutator that $[u, v]|_N = [u|_N, v|_N]$. Therefore $[u, v]$ is tangent to N.

1.5. The commutator of vector fields has three algebraic properties: it is skew-symmetric, bilinear and satisfies the following *Jacobi identity*.

Lemma. *For every three vector fields one has*

$$[[u, v], w] + [[v, w], u] + [[w, u], v] = 0. \tag{2}$$

Proof. One has:

$$L_{[[u,v],w]} = L_{[u,v]}L_w - L_w L_{[u,v]} = (L_u L_v - L_v L_u)L_w - L_w(L_v L_u - L_u L_v).$$

The sum (2) contains 3 such blocks, 12 terms altogether, and they cancel pairwise.

Definition. A linear space with a skew-symmetric bilinear operation (commutator) satisfying the Jacobi identity (2) is called *Lie algebra*. An isomorphism of Lie algebras is a linear isomorphism that takes commutator to commutator.

Example. One example of Lie algebras is known from Calculus 3: it is the algebra of vectors in 3-space with respect to the cross-product.

We have just proved that an associative algebra can be made into a Lie algebra by setting: $[A, B] = AB - BA$. For example, starting with the algebra of $n \times n$ matrices, real or complex, one arrives at the Lie algebras $gl(n, \mathbf{R})$ and $gl(n, \mathbf{C})$, respectively.

Exercise. Consider the associative algebra of 2×2 real matrices with zero trace and form the Lie algebra as above: $[A, B] = AB - BA$; this Lie algebra is called $sl(2, \mathbf{R})$. Prove that $sl(2, \mathbf{R})$ is isomorphic to the Lie algebra of vector fields on the line $\partial_x, x\partial_x, x^2\partial_x$ and that $sl(2, \mathbf{R})$ is not isomorphic to the Lie algebra of vectors in 3-space with respect to the cross-product.

1.6. Recall the following definition (from the previous topics course, Differential Topology).

Definition. A Lie group is a smooth manifold G which is also a group, and the two structures agree: the inversion map $G \to G$ and the multiplication map $G \times G \to G$ are smooth.

Examples. $\mathbf{R}^n$ is a group with respect to the vector summation, and so is the torus $T^n = \mathbf{R}^n/\mathbf{Z}^n$; these groups are commutative. $GL(n, \mathbf{R})$ is the group of non-degenerate $n \times n$ matrices; $SL(n, \mathbf{R})$ is its subgroup consisting of the matrices with determinant 1; $O(n) \subset GL(n, \mathbf{R})$ consists of the matrices that preserve a fixed scalar product (equivalently, $AA^* = E$ where E is the unit matrix); $SO(n) \subset O(n)$ consists of orientation preserving matrices. All these groups have complex versions; $U(n) \subset GL(n, \mathbf{C})$ consists of the matrices that preserve a fixed Hermitian product (equivalently, $AA^* = E$ where A^* is the transpose complex conjugated matrix), and $SU(n) \subset U(n)$ consists of the matrices with unit determinant.

Exercise. $SO(2)$ is diffeomorphic to the circle, and so is $U(1)$. The group $SU(2)$ is diffeomorphic to S^3.

It is also well known that $SO(3) = \mathbf{RP}^3$ (see Differential Topology or ask those who attended).

Given a Lie group G let $g = T_eG$ be the tangent space at the unit element. Similarly to Section 1.3, one makes g into Lie algebra. Vector $v, u \in g$ determine 1-parameter subgroups $\phi_t, \psi_s \subset G$. Then there exists a unique vector $w \in g$ such that the respective 1-parameter subgroup $\eta_\tau \subset G$ satisfies the following property:

$$\eta_{st} = \phi_t \psi_s \phi_{-t} \psi_{-s} \mod o(s^2 + t^2).$$

This vector w is called the commutator: $w = [v, u]$. For example, if G is a commutative group then g is a trivial Lie algebra.

Of course, a smooth homomorphism of Lie groups induces a homomorphism of the respective Lie algebras, that is, a linear map that takes commutator to commutator.

The explanation of the above construction is as follows. Every $x \in G$ determines a diffeomorphism $R_x : G \to G$ given by $R_x(y) = yx$. Given $v \in g$ we obtain a right-invariant vector field whose value at x is $dR_x(v) \in T_xG$. Thus we embed g in the space of vector fields on G, and g inherits the commutator from this Lie algebra.

On the other hand, the Lie algebra of vector fields on a manifold is the Lie algebra of the "Lie group" of diffeomorphisms of this manifold (caution is to be exercised since everything is infinite-dimensional here).

Example. Let $G = GL(n, \mathbf{R})$. Then g is the space of matrices. Given a matrix A, the respective 1-parameter subgroup is $\exp(tA)$. One has:

$$e^{tA} e^{tB} e^{-tA} e^{-tB} = E + st(AB - BA) + o(s^2 + t^2) = e^{st[A,B]},$$

therefore the Lie algebra structure in g is given by the commutator $[A, B] = AB - BA$. This Lie algebra is denoted by $gl(n, \mathbf{R})$. The same formula defines the commutator in other matrix Lie algebras $sl(n, \mathbf{R}), o(n), so(n), u(n), su(n)$, etc.

Exercise. Prove that $sl(n)$ consists of the traceless matrices and $o(n)$ of the matrices satisfying $A^* = -A$.

Assume that a Lie group G acts on a manifold M; this means that one has a smooth homomorphism from G to the group of diffeomorphisms of M. This homomorphism induces a homomorphisms of Lie algebras $g \to Vect(M)$, the Lie algebra of vector fields on M. Thus g has a *representation* in $Vect(M)$.

Example. The group $SL(2, \mathbf{R})$ acts on the real projective line by fractional-linear transformations. One obtains a homomorphism from the Lie algebra $sl(2, \mathbf{R})$ to the Lie algebra of vector fields on $\mathbf{RP}^1$. Choosing a coordinate x on $\mathbf{RP}^1$, the image of $sl(2, \mathbf{R})$ consists of the fields $\partial_x, x\partial_x, x^2\partial_x$.

Let $G_1 \to G_2$ be a covering of Lie groups, for example, $\mathbf{R}^n \to T^n$ or $SU(2) = S^3 \to \mathbf{RP}^3 = SO(3)$. Since a covering is a local diffeomorphism and the construction of Lie algebra is infinitesimal, the respective Lie algebras coincide: $g_1 = g_2$. In particular, $su(2) = so(3)$. Note also that $so(n) = o(n)$ since $O(n)$ consists of two components, and only the component of the unit element, $SO(n)$, is involved in the construction of the Lie algebra.

The theory of Lie algebras and Lie groups is very much developed, up to strong classification results. One of the results is that if g is the Lie algebra corresponding to a compact Lie group G then G is determined by g up to a covering. We will mention only one notion of this theory.

Definition. Given a Lie algebra g, a *Killing metric* is a metric satisfying

$$([x,y],z) + (y,[x,z]) = 0$$

for all $x, y, z \in g$.

Lemma. *The formula* $(X,Y) = -Tr(XY)$ *defines a Killing metric on the Lie algebra* $so(n)$.

Proof. If $X = x_{ij}$ then $(X,X) = \sum x_{ij}^2$, the usual Euclidean structure in space of dimension n^2. Its restriction to $so(n)$ is a Euclidean structure too. One needs to check that

$$Tr((XY - YX)Z + Y(XZ - ZX)) = 0$$

for $X, Y, Z \in so(n)$. This equals

$$Tr(XYZ - YXZ + YXZ - YZX) = Tr(XYZ - YZX) = 0$$

since the trace is invariant under cyclic permutations.

Likewise, one defines a Killing metric on $u(n)$ as Re $(-Tr(XY))$.

2. Differential forms.

2.1. Start with linear algebra. Let V^n be a linear space. We will deal with tensor powers $V \otimes ... \otimes V$ and exterior powers $\wedge^k V = V^{\otimes k}$ mod $(u \otimes v = -v \otimes u)$.

Question: what's the dimension of $\wedge^k V$?

A linear 1-form on a vector space V is a covector, i.e., an element of V^*. A linear 2-form (or an exterior 2-form) is a skew symmetric bilinear function ω on V:

$$\omega(u,v) = -\omega(v,u).$$

Example. If V is the plane then det (u,v) is a 2-form. More generally, consider n-dimensional space and fix a projection onto a 2-plane. Then the oriented area of the projection of a parallelogram is a 2-form. This example is most general.

Question: what's the dimension of the space of 2-forms? What is its relation to $\wedge^2 V^*$?

Likewise, a k-form on V^n is a skew symmetric k-linear function, that is, an element of $\wedge^k V^*$. For example, an n-form is proportional to the determinant of n vectors.

Exterior forms make an algebra with respect to the operation of exterior or wedge product.

Definition. Let α be a k-form and β an l-form. Define $\alpha \wedge \beta$ as the $k+l$-form whose value on vectors $v_1, ..., v_{k+l}$ equals

$$\sum (-1)^\mu \alpha(v_{i_1}, ..., v_{i_k}) \beta(v_{j_1}, ..., v_{j_l}),$$

where $i_1 < ... < i_k,\ j_1 < ... < j_l$, sum over permutations $(i_1, ..., i_k, j_1, ..., j_l)$ of $(1, ..., k+l)$, and μ is the sign of this permutation.

Lemma. *The exterior product is associative and skew-commutative:* $\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha$.

Exercise. Prove this lemma.

Example. If α is a k-form with k odd then $\alpha \wedge \alpha = 0$. On the other hand, consider $2n$-dimensional space with coordinates $p_i, q_i, i = 1, ..., n$, and let $\omega = \sum p_i \wedge q_i$. Then $\omega \wedge ... \wedge \omega$ (n times) is a non-zero $2n$-form.

Exercise. Consider $\mathbf{R}^3$. Given a vector v, consider the 1-form α_v and the 2-form ω_v defined by

$$\alpha_v(u) = (u, v); \quad \omega_v(u, w) = \det\ (v, u, w).$$

Prove the following:

$$\alpha_v \wedge \alpha_u = \omega_{v \times u}; \quad \alpha_v \wedge \omega_u = (u, v)\ \det.$$

This exercise shows that the cross-product is a particular case of the exterior product of forms.

Consider two linear spaces U and V and a linear map $f : U \to V$. Given a k-form α on V, one has a k-form $f^*(\alpha)$ on U: its value on vectors $u_1, ..., u_k$ is $\alpha(f(u_1), ..., f(u_k))$. This correspondence enjoys an obvious property $(fg)^* = g^* f^*$.

2.2. Let M be a smooth manifold.

Definition. A differential 1-form on M is a smooth function $\alpha(v, x)$ where $x \in M$ and $v \in T_x M$; for every fixed x this function is a linear 1-form on the tangent space $T_x M$.

Examples. If f is smooth function on M then df is a differential 1-form. Not every differential 1-form is the differential of a function: if x is the angular coordinate on the circle then dx is a differential 1-form which fails to be the differential of a function.

Differential 1-forms in $\mathbf{R}^n$ can be described as follows. Choose coordinates $x_1, ..., x_n$. Then every differential 1-form is written as $f_1 dx_1 + ... + f_n dx_n$ where $f_1, ..., f_n$ are smooth functions of $x_1, ..., x_n$.

Definition. A differential k-form on M is a smooth function $\alpha(v_1, ..., v_k, x)$ where $x \in M$ and $v_i \in T_x M,\ i = 1, ..., k$; for every fixed x this function is a linear k-form on the tangent space $T_x M$.

Differential forms form a vector space, and they can be multiplied by smooth functions. The exterior product of differential p and q forms is a a differential $p+q$-form. A differential k-form in $\mathbf{R}^n$ can be written as

$$\sum_{i_1 < ... < i_k} f_{i_1, ..., i_k} dx_{i_1} \wedge ... \wedge dx_{i_k}.$$

Since every manifold M^n is locally diffeomorphic to $\mathbf{R}^n$, one has a similar expression in local coordinates on M. Of course, there are no non-trivial k-forms for $k > n$.

To change coordinates in differential forms one uses the Important Exercise from Section 1.1.

Exercise. Express the forms $xdy - ydx$ and $dx \wedge dy$ in polar coordinates.

2.3. Differential forms are made to integrate (recall Calculus 3). Start with a particular case. Let $P \subset \mathbf{R}^n$ be a convex polyhedron and $\omega = f(x)dx_1 \wedge ... \wedge dx_n$ an n-form. Define

$$\int_P \omega = \int_P f(x)dx_1...dx_n.$$

Clearly, this integral is linear in ω. If P is decomposed into two polyhedra P_1 and P_2 then

$$\int_P \omega = \int_{P_1} \omega + \int_{P_2} \omega.$$

Important Example. Consider two polyhedra P_1 and P_2 in $\mathbf{R}^n$ and let $f : P_1 \to P_2$ be an orientation-preserving diffeomorphism. Then for every n-form one has:

$$\int_{P_1} f^*(\omega) = \int_{P_2} \omega.$$

Indeed, the RHS is

$$\int_{P_2} f(y)dy_1...dy_n = \int_{P_1} \frac{D(y)}{D(x)} f(y(x))dx_1...dx_n = \int_{P_1} f^*(\omega),$$

the first equality being the change of variables formula and the second following from the definition of $f^*(\omega)$.

In general, one integrates a k-form on M^n over a k-chain. The definition will generalize the integral of a 1-form over a curve (recall Cal 3 again).

Definition. A singular k-dimensional polyhedron σ is an oriented convex polyhedron $P \subset \mathbf{R}^k$ and a smooth map $f : P \to M$. Then $-\sigma$ differs from σ by the orientation of $\mathbf{R}^k$. A k-chain is a linear combination of singular k-dimensional polyhedra. Given a k-form ω on M, one defines

$$\int_\sigma \omega = \int_P f^*(\omega).$$

The definition of integral is extended to chains by linearity: if $c = \sum k_i \sigma_i$ then

$$\int_c \omega = \sum k_i \int_{\sigma_i} \omega.$$

Exercises. 1. Show that $\int_{-\sigma} \omega = -\int_\sigma \omega$.

2. Let $f : M^n \to N^n$ be a k-fold covering of compact manifolds and ω is an n-form on N. Prove that

$$\int_M f^*(\omega) = k \int_N \omega.$$

2.4. Let $P \subset \mathbf{R}^n$ be an oriented convex polyhedron.

Definition. The boundary ∂P is the $n-1$-chain whose singular polyhedra are the faces of P oriented by the outward normals. The boundary of a singular polyhedron is defined analogously and the definition is extended to chains by linearity.

Lemma. $\partial^2 = 0$.

Proof. One needs to check that each codimension 2 face of P appears in $\partial^2(P)$ twice with opposite signs. Intersecting by a plane, the claim reduces to the case of polygons which is obvious.

2.5. The familiar differentiation operation $f \to df$ extends to an operation of *exterior differentiation* that assigns a differential $k+1$-form $d\omega$ to a differential k-form ω.

Given tangent vectors $v_1, ..., v_{k+1} \in T_x M^n$, we define $d\omega(v_1, ..., v_{k+1})$ as follows. Choose a coordinate system near point x; such a choice identifies $v_1, ..., v_{k+1}$ with vectors in $T_0\mathbf{R}^n = \mathbf{R}^n$. Let P be the parallelepiped generated by these vectors. The coordinate system provides a singular polyhedron $P \to M$. Let

$$L(v_1, ..., v_{k+1}) = \int_{\partial P} \omega.$$

Definition-Proposition. *The principal $k+1$-linear part of $L(v_1, ..., v_{k+1})$ is a skew-symmetric linear function, independent of the choice of the coordinate system:*

$$d\omega(v_1, ..., v_{k+1}) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^{k+1}} L(\varepsilon v_1, ..., \varepsilon v_{k+1}).$$

If, in local coordinates,

$$\omega = \sum f_{i_1,...,i_k} dx_{i_1} \wedge ... \wedge dx_{i_k}$$

then

$$d\omega = \sum df_{i_1,...,i_k} \wedge dx_{i_1} \wedge ... \wedge dx_{i_k}.$$

Example. Consider the case of 0-forms, i.e., functions. Let f be a function on M. The singular polyhedron is a curve on M from point x to point y whose tangent vector is $v \in T_x M$. Then $L(v) = f(y) - f(x)$, and the principal linear part of this increment is $df(v)$.

Sketch of Proof. Let us make a computation is a particular case: $\omega = f(x,y)dx$. Let u and v be two vectors (which will be multiplied by ε). The sides of the parallelogram P are given by:

$$t \to tu, \ t \to u + tv, \ t \to tv, \ t \to v + tu; \ t \in [0,1].$$

Therefore

$$\int_{\partial P} \omega = \int_0^1 (f(tu) - f(v+tu))u_1 - (f(tv) - f(u+tv))v_1 dt.$$

Next,

$$f(v+tu) - f(tu) = f_x(0,0)v_1 + f_y(0,0)v_2 + (\varepsilon^2)$$

and

$$f(u+tv) - f(tv) = f_x u_1 + f_y u_2 + (\varepsilon^2).$$

Thus

$$\int_{\partial P} \omega = f_y(u_2 v_1 - u_1 v_2) + (\varepsilon^3),$$

that is,

$$d\omega = f_y dy \wedge dx = df \wedge dx.$$

If one changes the coordinate system then the curvilinear parallelogram P will be replaced by a new one, P'; however

$$\int_{\partial P'} \omega - \int_{\partial P} \omega$$

is cubic in ε. The general case is similar but more cumbersome to compute.

2.6. The exterior differentiation enjoys the following properties.

Lemma. *(i)* $d(\omega + \eta) = d\omega + d\eta$.
(ii) $d(\omega^k \wedge \eta^l) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta$.
(iii) If $f : M \to N$ *is a smooth map and* ω *is a form on* N *then* $df^*(\omega) = f^* d(\omega)$.
(iv) $d^2\omega = 0$.

Proof. The first is obvious, the third immediately follows from the definition. The second is best checked in local coordinates:

$$\omega = \sum f_I dx_{i_1} \wedge ... \wedge dx_{i_k}, \ \eta = \sum g_J dx_{j_1} \wedge ... \wedge dx_{j_l}.$$

The last property is also verified in local coordinates.

2.7. We are ready to prove an important result.

Theorem. *Given a* $k+1$*-chain* c *and a* k*-form* ω*, one has*

$$\int_{\partial c} \omega = \int_c d\omega.$$

Proof. Consider first the case when c consists of one singular cube $f : P = I^{k+1} \to M$. Partition P into N^{k+1} equal cubes P_i. Let $\varepsilon = 1/n$. Consider a small cube P_i with edges $v_1^i, ..., v_{k+1}^i$ of order ε. Then, according to the definition,

$$d\omega(v_1^i, ..., v_{k+1}^i) = \int_{\partial P_i} \omega + o(\varepsilon^{(k+1)}).$$

It follows that

$$\sum_{i=1}^{N^{k+1}} d\omega(v_1^i, ..., v_{k+1}^i) = \int_{\partial P} \omega + o(\varepsilon).$$

The LHS being the Riemann sum for

$$\int_P d\omega,$$

the result follows by taking the limit $\varepsilon \to 0$.

Next, one proves the formula for a simplex; the general result will follow since every polyhedron partitions into simplices. The result for a simplex follows from that for a cube: there is a smooth map from a cube to a simplex which is an orientation preserving diffeomorphism in the interior and which is an orientation preserving diffeomorphism on some faces while other faces are sent to faces of smaller dimensions. Alternatively, one may approximate a polyhedron by a union of cubes; we do not elaborate.

Examples. 1. Let us deduce the classical Green theorem. Let P be an oriented polygonal domain in the plane, and $\omega = fdx + gdy$ is a 1-form. Then $d\omega = (g_x - f_y)dx \wedge dy$, and we have:

$$\int_{\partial P} fdx + gdy = \int_P (g_x - f_y)dxdy.$$

2. Next we deduce the divergence theorem

$$\int_P divFdV = \int_{\partial P} (F \cdot n)dA$$

where P is an oriented domain in 3-space. Let

$$\omega = F_1 dy \wedge dz + F_2 dz \wedge dx + F_3 dx \wedge dy.$$

Then $d\omega = divFdV$. On the other hand, we remember from Cal 3 (if not, prove it!) that

$$n_1 dA = dy \wedge dz,\ n_2 dA = dz \wedge dx,\ n_3 dA = dx \wedge dy.$$

Therefore $(F \cdot n)dA = \omega$, and the divergence theorem follows.

3. Finally, we deduce the Stokes theorem. Let P be an oriented surface in 3-space with boundary. Let T be the unit tangent vector field along ∂P and F a vector field along P. Then

$$\int_P (curlF \cdot n)dA = \int_{\partial P} (F \cdot T)ds.$$

Consider the 1-form $\omega = F_1 dx + F_2 dy + F_3 dz$. Then, as above, one has:

$$(curlF \cdot n)dA = d\omega.$$

On the other hand,

$$T_1 ds = dx,\ T_2 ds = dy,\ T_3 ds = dz,$$

so $(F \cdot T)ds = \omega$. The Stokes theorem follows.

2.8. Since $d^2\omega = 0$ one has: Im $d \subset$ Ker d. One defines the *de Rham cohomology* of a smooth manifold:

$$H^i_{DR}(M) = \text{Ker } d/\text{Im } d,$$

where d is taken on i-forms on M. A form ω is called *closed* if $d\omega = 0$; a form ω is called *exact* if $\omega = d\eta$. Thus i-th de Rham cohomology is the quotient space of closed i-forms by the exact ones.

Examples. 1. Let M be a connected manifold. Then $H^0_{DR}(M) = \mathbf{R}$. Indeed, if $df = 0$ then f is constant.

2. $H^1_{DR}(\mathbf{R}) = 0$ since for every $f(x)$ one has: $fdx = dg$ where $g = \int f$. In contrast, $H^1_{DR}(S^1) = \mathbf{R}$ since fdx is exact only if $\int fdx = 0$.

3. Let M^n be a closed oriented manifold. Then $H^n_{DR}(M) = \mathbf{R}$. Indeed, every n-form is closed, and for every $n-1$-form ω one has:

$$\int_M d\omega = \int_{\partial M} \omega = 0.$$

Exercise. Compute $H^1_{DR}(T^2)$.

Integration provides a pairing between chains and forms; Theorem 2.7 implies that one also has a pairing between singular homology (with real coefficients) and de Rham cohomology. De Rham's theorem asserts that this pairing is non-degenerate:

$$H^*_{DR}(M) = H_*(M, \mathbf{R}).$$

For example, the statement that

$$H^i_{DR}(\mathbf{R}^n) = 0,\ i > 0$$

is called the Poincaré lemma. A proof can be deduced from Theorem 2.7. Let P be a singular polyhedron in $\mathbf{R}^n$. Denote by CP the cone over P. Then one has:

$$\partial CP + C\partial P = P.$$

In particular, if $\partial c = 0$ then $c = \partial b$. Given a form ω, define $H\omega$ by

$$\int_c H\omega = \int_{C(c)} \omega$$

for every chain c. Then $dH + Hd = id$; one can write an explicit formula for $H\omega$. It follows that Ker d = Im d.

One of the advantages of de Rham cohomology is that the ring structure is very transparent: the multiplication is induced by the wedge product of differential forms.

Exercise. Prove that the wedge product induces a well-defined multiplication of de Rham cohomology classes.

2.9. Let us discuss relations between vector fields and differential forms. Let ω be a k-form and v a vector field. Define a $k-1$-form $i_v\omega$ by the formula:

$$i_v\omega(u_1, ..., u_{k-1}) = \omega(v, u_1, ..., u_{k-1}).$$

Exercise. Show that $i_v i_u \omega = -i_u i_v \omega$.

Next, define the *Lie derivative*. Let v be a vector field and ω a differential k-form. Let ϕ_t be the respective 1-parameter group of diffeomorphisms. Then $L_v\omega$ is a k-form such that for every chain c one has:

$$\int_c L_v\omega = \frac{d}{dt}|_{t=0} \int_{\phi_t(c)} \omega.$$

In particular, $L_v f = df(v)$ is the directional derivative.

Lemma (homotopy formula). *One has:*

$$i_v d\omega + d i_v \omega = L_v \omega.$$

Proof. Define the homotopy operator H: given a singular polyhedron $f : P \to M$, let $Hf : P \times [0,1] \to M$ be given by the formula $(x,t) \to \phi_t(x)$. Then one has:

$$\phi_1(c) - c = \partial Hc + H\partial c.$$

This implies the desired formula.

Exercise. Show that $dL_v = L_v d$.

Lemma (Cartan formula). *Let ω be a 1-form and u, v be vector fields. Then*

$$d\omega(u,v) = L_u(\omega(v)) - L_v(\omega(u)) - \omega([u,v]).$$

Proof. A direct computation in coordinates.

The following monster is the Cartan formula for k-forms:

$$d\omega(v_1, ..., v_{k+1}) = \sum (-1)^{i-1} L_{v_i} \omega(v_1, ..., \hat{v}_i, ..., v_{k+1}) +$$

$$\sum_{i<j} (-1)^{i+j} \omega([v_i, v_j], v_1, ..., \hat{v}_i, ..., \hat{v}_j, ..., v_{k+1});$$

the "hat" means that the respective term is omitted.

Exercise. Show that
1). $[L_u, i_v] = i_{[u,v]}$;
2). $[L_u, L_v] = L_{[u,v]}$.

Example. Let M^n be a manifold with a volume n-form μ, and v a vector field on M. Then $L_v \mu = f\mu$ where the function f is called the *divergence* of v. A computation in coordinates yields the familiar formula for the divergence.

2.10. Let M be a smooth manifold.

Definition. A k-dimensional *distribution* (or a field of k-planes) is given if at every point $x \in M$ a k-dimensional subspace $E^k \subset T_x M$ is given, smoothly depending on the point x. A distribution is oriented or cooriented if so is the k-space E at every point.

Example. A 1-dimensional oriented distribution E determined a vector field tangent to E. Conversely, every non-vanishing vector field integrates to a 1-dimensional oriented distribution.

Definition. A k-dimensional *foliation* is a partition of M^n into disjoint union of k-dimensional submanifolds locally diffeomorphic to the partition of $\mathbf{R}^n$ into k-dimensional parallel subspaces. The respective k-dimensional submanifolds are called the *leaves* of the foliation

Examples. 1. Start with a partition of $\mathbf{R}^n$ into k-dimensional parallel subspaces, and factorize by the integer lattice. One obtains a k-dimensional foliation on the torus.

2. Consider the foliation of the strip $|x| \leq \pi/2$ by the graphs $y = \tan^2 x + const$. Rotate the strip about the vertical axis to obtain a foliation of a solid cylinder. Factorize by a parallel translation in the vertical direction to obtain a foliation of a solid torus; the boundary torus is a leaf. The 3-sphere is made of two solid tori; we obtain a foliation of S^3 called the Reeb foliation.

A k-dimensional foliation determines a k-dimensional distribution. Is the converse true? Given a distribution E^k, consider the space of vector fields tangent to E and denote it by $V(E)$. Similarly, consider the space of differential 1-forms that vanish on E; denote by $\Omega(E)$.

Theorem (Frobenius). *A distribution E is a foliation iff $V(E)$ is a Lie algebra or, equivalently, $d\Omega(E) \subset \Omega(E) \wedge \Omega(M)$.*

Proof. If E is a foliation then, as we noticed before, one has: $[u, v] \in V(E)$ for all $u, v \in V(E)$.

Conversely, let $V(E)$ be a Lie algebra. Choose k linearly independent vector fields $v_1, ..., v_k$ in a small neighborhood that span E, and choose coordinates so that $v_1 = \partial x_1$. Let $f_i = dx_1(v_i)$, $i = 2, ..., k$, and set: $u_i = v_i - f_i v_1$. That is, u is obtained from v by deleting the term with $\partial/\partial x_1$. Then $dx_1(u_i) = 0$. Consider the family of hyperplanes $x_1 = const$. It follows that the fields u_i, $i = 2, ..., k$ are tangent to these hyperplanes. Moreover, that space generated by the fields u_i is a Lie algebra. By induction, we have a $k-1$-dimensional foliation F. The products of the leaves of this foliation and segments of the x_1-axis are the leaves of the desired foliation.

Choose a basis of vector fields $v_1, ..., v_n$ so that $v_1, ..., v_k$ generate E. Let $\alpha_1, ..., \alpha_n$ be the dual basis of 1-forms: $\alpha_i(v_j) = \delta_{ij}$. Then $\Omega(E)$ is generated (over smooth functions) by $\alpha_{k+1}, ..., \alpha_n$. According to the Cartan formula,

$$d\alpha_q(v_i, v_j) = L_{v_i}\alpha_q(v_j) - L_{v_j}\alpha_q(v_i) - \alpha_q([v_i, v_j]).$$

For $i, j \leq k$ and $q > k$ the first two terms on the RHS vanish. The RHS vanishes for all $q > k$ iff $[v_i, v_j]$ is a combination of $v_1, ..., v_k$, i.e., when $V(E)$ is a Lie algebra. The LHS vanishes for all $i, j \leq k$ iff $d\alpha_q$ belongs to the ideal generated by $\alpha_{k+1}, ..., \alpha_n$, that is, $d\Omega(E) \subset \Omega(E) \wedge \Omega(M)$.

Example-Exercise. Every codimension 1 distribution is locally given by a 1-form α. This distribution is a foliation iff $d\alpha = \alpha \wedge \beta$. Prove that this is equivalent to $\alpha \wedge d\alpha = 0$.

Exercise: Godbillion-Vey class. Let F be a cooriented codimension 1 foliation on a manifold M. Choose a 1-form $\alpha \in \Omega(F)$; then $d\alpha = \alpha \wedge \beta$ for some 1-form β. Prove that $\beta \wedge d\beta$ is a closed 3-form, and that its cohomology class does not depend on the choices involved.

Example. In 3-space with coordinates x, y, z consider the 1-form $\alpha = dz + y dx$. Then $\alpha \wedge d\alpha$ is a volume form. This 2-distribution E is an example of a *contact structure.* In terms of vector fields, ∂y and $y\partial z - \partial x$ form a basis of $V(E)$.

Exercise. Consider the space $\mathbf{C}^2$ and the unit sphere S^3 in it. For $x \in S^3$ define a 2-dimensional tangent space: $E^2(x) = T_x S^3 \cap \sqrt{-1} T_x S^3$ (the unique copy of $\mathbf{C}$ contained in the tangent space $T_x S^3$). Prove that E is a contact structure.

3. Symplectic geometry.

3.1. Start with symplectic linear algebra.

Definition. A symplectic structure ω in a linear space V is a non-degenerate skew-symmetric bilinear form.

Note that $\omega(u,u)=0$ for all $u \in V$.

Examples. An area form in the plane is a symplectic structure; one can choose a basis (p,q) such that $\omega(p,q)=1$. Taking a direct sum of a number of planes, one obtains a symplectic structure in every even-dimensional space with a basis $(p_1,...,p_n,q_1,...,q_n)$ such that the only non-trivial products are $\omega(p_i,q_i)=1,\ i=1,...,n$. Such a basis is called a *Darboux basis.*

Let U be a linear space. Then the space $U \oplus U^*$ has a symplectic structure defined as follows:

$$\omega(u_1,u_2)=\omega(l_1,l_2)=0 \ \text{ and } \ \omega(u,l)=l(u).$$

Exercise. Prove that this is a symplectic structure.

Lemma. *A symplectic space has an even dimension.*

Proof. Choose a Euclidean structure in V^n. Then ω is given by a linear operator A:

$$\omega(u,v)=(Au,v).$$

Since ω is skew-symmetric, one has: $A^*=-A$. Therefore $\det A=\det A^*=(-1)^n \det A$. If A is non-degenerate n must be even.

Similarly to a Euclidean structure, a symplectic structure identifies the space with its dual: to a vector $u \in V$ there corresponds the covector $\omega(u,\cdot) \in V^*$.

Theorem (linear Darboux theorem). *All $2n$-dimensional symplectic spaces are linearly symplectically isomorphic.*

Proof. Induction in n. For $n=1$ the claim is obvious. Choose a pair of vectors $p_1,q_1 \in V^{2n}$ such that $\omega(p_1,q_1)=1$, and let U be the plane spanned by these vectors. Consider the orthogonal complement to U with respect to ω; this is a $2n-2$-dimensional space, say, W. Then W is a symplectic space. By the induction assumption, it has a basis $(p_2,...,p_n,q_2,...,q_n)$ as in the above example. One adds (p_1,q_1) to this basis, and the result follows.

Definition. A subspace $L \subset V^{2n}$ of a symplectic space is called *Lagrangian* if ω vanishes on L and L has the maximal possible dimension.

Lemma. *This dimension is equal to n.*

Proof. Let L' be the orthogonal complement of L. Then $L \subset L'$ and dim $L'=2n-$ dim L. Therefore dim $L \le n$. Examples of Lagrangian subspaces are provided by the p- or q-spaces in a Darboux basis.

Clearly, every 1-dimensional subspace of the plane is Lagrangian.

Exercise. Consider the symplectic space $U \oplus U^*$. Let $A : U \to U^*$ be a linear map. Prove that the graph of A is a Lagrangian subspace iff $A^* = A$. Conclude that the manifold of all Lagrangian subspaces in $2n$-dimensional symplectic space has dimension $n(n+1)/2$.

3.2. **Definition**. A *symplectic structure* on a manifold M is a non-degenerate closed differential 2-form ω. Given two symplectic manifolds (M_1, ω_1) and (M_2, ω_2), a smooth map $f : M_1 \to M_2$ is *symplectic* if $f^*(\omega_2) = (\omega_1)$. A symplectic diffeomorphism is called a *symplectomorphism.*

The tangent space of a symplectic manifold is a linear symplectic space; in particular, dim M is even. The condition that ω is closed is harder to "visualize".

Examples 1. Let M be $\mathbf{R}^{2n}$ with Darboux coordinates $(p_1, ..., p_n, q_1, ..., q_n)$. Then $\omega = dp_1 \wedge dq_1 + dp_2 \wedge dq_2 + ... + dp_n \wedge dq_n$ is a symplectic structure. The same formula defines a symplectic structure on $2n$-dimensional torus (p and q are cyclic coordinates). The natural embedding $j : \mathbf{R}^{2n} \subset \mathbf{R}^{2n+2}$ is symplectic.

2. Let M be a smooth manifold. Then the cotangent bundle T^*M has a canonical symplectic structure ω. First, one defines a canonical 1-form λ called the Liouville form. Let $p \in T_x^*M$ and let v be a tangent vector to T^*M at point p. Consider the projection $\pi : T^*M \to M$; then $d\pi(v) \in T_xM$. We define $\lambda(v)$ as $p(d\pi(v))$. Then one defines $\omega = d\lambda$. A diffeomorphism of M induces a symplectomorphism of T^*M.

To express λ is coordinates let $(q_1, ..., q_n)$ be local coordinates on M. Then $(p_1 = dq_1, ..., p_n = dq_n)$ is a basis in every cotangent space T_x^*M. Thus $(p_1, ..., p_n, q_1, ..., q_n)$ are local coordinates in T^*M, and one has: $\lambda = p_1dq_1 + ... + p_ndq_n$. Therefore ω is a symplectic form.

Cotangent bundles appear in classical mechanics as phase spaces of mechanical systems; we will later discuss this in some detail.

3. Let M be a surface. A symplectic structure on M is just an area form (it is closed automatically). Thus every oriented surface has a symplectic structure. A symplectic map of surfaces is simply an orientation and area-preserving map.

Note that if (M, ω) is a closed (compact, no boundary) symplectic manifold then ω defines a non-trivial de Rham cohomology class. For example, the sphere S^{2n} is not a symplectic manifold unless $n = 1$.

Exercise (Archimedes). Consider a sphere in 3-space and a circumscribed cylinder. Let f be the radial projection from the sphere to the cylinder. Prove that f is a symplectic map.

3.3. Euclidean structure identifies tangent and cotangent spaces and associates the gradient vector field with a function. Similarly, a symplectic structure on a manifold M associates the *symplectic gradient* sgrad H with a function $H : M \to \mathbf{R}$ (the function H is often called a Hamiltonian and the vector field sgrad H a Hamiltonian vector field).

Definition. The field sgrad H is defined by the equality $\omega(v, \text{sgrad } H) = dH(v)$ that holds for every vector field v on M. Equivalently, $i_{\text{sgrad } H}\omega = -dH$.

In other words, sgrad H is the vector field dual to the 1-form dH with respect to the symplectic structure ω.

Example. Compute sgrad H in Darboux coordinates in which $\omega = \sum dp_i \wedge dq_i$. Let sgrad $H = \sum a_i\partial p_i + b_i\partial q_i$. One has: $i_{\text{sgrad } H}\omega = \sum a_idq_i - b_idp_i$. On the other hand,

$dH = \sum H_{q_i} dq_i + H_{p_i} dp_i$. Thus $a_i = -H_{q_i}, b_i = H_{p_i}$, and

$$\text{sgrad } H = \sum H_{p_i} \partial q_i - H_{q_i} \partial p_i.$$

Exercises. 1. Prove that H is constant along the trajectories of the vector field sgrad H (hint: find the directional derivative of H along this field).

2. Find the integral curves of the fields sgrad $x^2 \pm y^2$ in the plane with the symplectic structure $dx \wedge dy$.

3. Compute the vector field sgrad z on the unit sphere $x^2 + y^2 + z^2 = 1$ with its standard area form.

Lemma. *A Hamiltonian vector field preserves the symplectic structure.*

Proof. Let $v = \text{sgrad } H$; we want to show that $L_v \omega = 0$. Indeed,

$$L_v \omega = i_v d\omega + d i_v \omega = -ddH = 0,$$

and we are done.

It follows that the respective 1-parameter group of diffeomorphisms consists of symplectomorphisms.

Conversely, if $L_v \omega = 0$ for some vector field v then $i_v \omega$ is closed, and therefore *locally* there exists a function H such that $i_v \omega = -dH$. Thus v is a locally Hamiltonian vector field. However the function may not exist globally: consider, for example, the field ∂_x on the torus with the symplectic structure $dx \wedge dy$.

3.4. One can define a Lie algebra structure on the space of functions on a symplectic manifold. Let (M, ω) be a symplectic manifold and f, g two smooth functions. Define the *Poisson bracket*

$$\{f, g\} = dg \text{ (sgrad } f).$$

In other words,

$$\{f, g\} = \omega \text{ (sgrad } f, \text{sgrad } g).$$

Example. Compute the Poisson bracket in Darboux coordinates:

$$\{f, g\} = \sum f_{p_i} g_{q_i} - f_{q_i} g_{p_i}.$$

Exercise. Show that the Poisson bracket satisfies the Leibnitz identity:

$$\{f_1 f_2, g\} = f_1 \{f_2, g\} + f_2 \{f_1, g\}.$$

Clearly, the Poisson bracket is a skew-symmetric bilinear operation on functions.

Proposition. *1. The Poisson bracket satisfies the Jacobi identity:*

$$\{\{f,g\},h\}+\{\{g,h\},f\}+\{\{h,f\},g\}=0.$$

2. One has:

$$[\text{sgrad } g, \text{sgrad } h] = \text{sgrad } \{g,h\}.$$

Proof. The sum in question is a combination of second partial derivatives of the three functions involved. Let u, v, w be the symplectic gradient of the functions f, g, h. The second partial derivatives of f appear as follows:

$$\{\{f,g\},h\}+\{\{h,f\},g\}=L_w L_v(f)-L_v L_w(f)=L_{[w,v]}(f).$$

The last expression involves only one derivative, thus the sum in question is free from second derivatives of f. The same applies to g and h, therefore, the sum is zero. This proves the first claim.

The first claim thus can be written as

$$L_{[w,v]}(f)+\{\{g,h\},f\}=0$$

or

$$L_{[v,w]}=L_{\text{sgrad } \{g,h\}}.$$

The second claim follows.

Therefore the mapping $f \to \text{sgrad } f$ is a Lie algebra homomorphism.

Exercise. Prove the above Proposition by a direct computation in Darboux coordinates.

3.5. Unlike Riemannian geometry, symplectic geometry does not have local invariants. More precisely, one has the next result.

Theorem (Darboux). *Every two symplectic manifolds of the same dimension are locally symplectomorphic.*

Proof. We may assume that two symplectic forms ω_0 and ω_1 are given in a neighborhood of the origin $U \subset \mathbf{R}^{2n}$. Moreover, since every two linear symplectic structures are linearly isomorphic, we may assume that the restrictions of both forms to the tangent space at the origin coinide. We will construct a diffeomorphism f of a possibly smaller neighborhood of the origin such that $f^*(\omega_1)=\omega_0$.

Use the following homotopy method. The form $\omega_t=(1-t)\omega_0+t\omega_1$ is a linear symplectic structure on the tangent space at the origin, and $d\omega_t=0$. Therefore ω_t is a symplectic structure in a small neighborhood of the origin. Moreover, $\omega_t=\omega_0+td\sigma$ where σ is some 1-form. We will find a family of local diffeomorphisms f_t, fixing the origin, and such that $f_t^*(\omega_t)=\omega_0$.

The main idea is to represent the maps f_t as the flow of a time-dependent vector field v_t:

$$\frac{df_t}{dt}(x)=v_t(f_t(x)).$$

Since $f_t^*(\omega_t) = \omega_0$, one has:

$$0 = \frac{df_t^* \omega_t}{dt} = f_t^*(L_{v_t}\omega_t + \frac{d\omega_t}{dt}) = f_t^*(L_{v_t}\omega_t + d\sigma).$$

Thus

$$L_{v_t}\omega_t + d\sigma = 0 \quad \text{or} \quad i_{v_t}\omega_t + d\sigma = 0.$$

Since ω_t is non-degenerate, the last equation uniquely determines the desired v_t, and we are done.

Darboux theorem implies that, without loss of generality, any local computation can be made in Darboux coordinates. Darboux theorem is only a local result: there may be many non-equivalent symplectic structures on the same manifold.

Example. Consider the plane with a linear symplectic form and an open disk. These are diffeomorphic but not symplectomorphic manifolds: the area of the former is infinite and that of the latter is finite.

3.6. Let (M^{2n}, ω) be a symplectic manifold and $N^{2n-1} \subset M$ a smooth hypersurface. Then the restriction of ω to N is not non-degenerate anymore: it has a 1-dimensional kernel at every point.

Definition. The *characteristic direction* $\xi_x \subset T_x N$ is Ker $\omega | T_x N$. The *characteristic foliation* $\mathcal{F}$ is the 1-dimensional foliation on N whose tangent line at every point is the characteristic direction.

Let N be a non-singular level surface of a smooth function H.

Lemma. *For every $x \in N$ one has: sgrad $H(x) = \xi_x$.*

Proof. One has: $i_{\text{sgrad } H}\omega = dH = 0$ on N.

Assume that the space of leaves of the characteristic foliation $\mathcal{F}$ is a smooth manifold X^{2n-2} (locally this is always the case).

Theorem. *X^{2n-2} has a canonical symplectic structure.*

Proof. Let $p : N \to X$ be the projection. Define a 2-form Ω on X as follows. Let u_1, u_2 be tangent vectors to X at point x. Choose a point $y = p^{-1}(x)$ and tangent vectors $v_1, v_2 \in T_y N$ such that $dp(v_i) = u_i,\ i = 1, 2$. Set: $\Omega(u_1, u_2) = \omega(v_1, v_2)$.

We need to show that this is well defined. One may change v_i by an element of Ker dp, that is, by a vector from ξ_y, but such a change does not effect $\omega(v_1, v_2)$. One may also change the point y. Such a change is induced by the flow of the vector field sgrad H from the preceding lemma. Since sgrad H preserves the form ω, this does not effect $\omega(v_1, v_2)$ either.

Since $TX = TN/\xi$ and ξ is the kernel of ω, the form Ω is non-degenerate. It remains to show that Ω is closed. Consider a point $x \in X$, and choose a point $y = p^{-1}(x)$. Consider a small $2n-2$-dimensional disk V that contains y and is transversal to $\mathcal{F}$. Then $U = p(U)$ is a neighborhood of x, and $p : V \to U$ is a diffeomorphism. Moreover, p takes Ω to $\omega|_V$. Since ω is closed, so is Ω.

The above construction is called *symplectic reduction.*

Two Important Examples. 1). Consider Euclidean space $\mathbf{R}^{n+1}$ and identify the tangent and cotangent spaces using the Euclidean structure. Then $T\mathbf{R}^{n+1} = \mathbf{R}^{2n+2} = T^*\mathbf{R}^{n+1}$ is a symplectic manifold with the Darboux symplectic structure $dp \wedge dq$, where (q,p) is a tangent vector with foot point q. Consider the function $H(q,p) = p^2/2$. Then sgrad $H = p\partial_q$; this is a constant flow with speed $|p|$. Consider the "unit energy" hypersurface $H = 1$, and make the symplectic reduction construction. The quotient space is the space $\mathcal{L}$ of oriented lines in $\mathbf{R}^{n+1}$. We conclude that $\mathcal{L}$ is a $2n$-dimensional symplectic manifold. Moreover, one has the following result.

Lemma. *The space $\mathcal{L}$ is symplectomorphic (up to the sign) to T^*S^n.*

Proof. Let l be an oriented line. Denote by p the unit vector in the direction of l and by q the foot of the perpendicular from the origin to l. The map $l \to (q,p)$ defines an embedding of $\mathcal{L}$ to $T\mathbf{R}^{n+1}$, transversal to the trajectories of the field sgrad H. Therefore the symplectic structure on $\mathcal{L}$ is $dp \wedge dq$. On the other hand, q is perpendicular to p, therefore the set of pairs (p,q) is TS^n (where p is a point of the sphere and q is a tangent vector). Identifying the tangent and cotangent spaces again, we obtain the result.

In particular, the space of oriented lines in the plane is the cylinder with the area form $dp \wedge d\alpha$. One can reconstruct the metric of the plane from this area form. Namely, the next result holds.

Lemma. *Let γ be a plane curve. Given an oriented line l, let $f(l)$ be the number of intersection points $l \cap \gamma$. Then*

$$\text{length } \gamma = \frac{1}{4} \int_{\mathcal{L}} f(l) dp \wedge d\alpha.$$

Proof. It suffices to consider a broken line. Since both sides of the equality are additive, it is enough to consider one segment. For a segment, this is a direct computation.

The above result belongs to *integral geometry*.

Exercises. 1. Prove that the space of non-oriented lines in the plane is diffeomorphic to an open Moebius band.

2. Repeat the construction from the previous example replacing Euclidean space by S^2. Show that the quotient space $\mathcal{L}$ is the space of oriented great circles in S^2, and that $\mathcal{L}$ is symplectomorphic to S^2 with its standard area form.

3. Start with an area form $\phi(p,\alpha) dp \wedge d\alpha$ on the space of oriented line sin the plane; here $\phi(-p, \alpha + \pi) = \phi(p, \alpha) > 0$ is a positive even function. The formula from the above lemma defines a new "length" in the plane. Prove that this length satisfies the triangle inequality, that is, the straight segment is the shortest curve between two points.

2). Clearly $\mathbf{C}^n$ is a symplectic manifold. Apply symplectic reduction to show that so is $\mathbf{CP}^n$. Start with $\mathbf{C}^{n+1}$ with its Darboux symplectic structure, and let $H(q,p) = (q^2 + p^2)$. Then sgrad $H = 2(p\partial_q - q\partial_p)$. The unit energy hypersurface is the unit sphere, and the vector field is the Hopf field $v(z) = iz$. The trajectories are $exp(it)z$, and the quotient space is $\mathbf{CP}^n$.

This example provides a link between symplectic and algebraic geometry: a smooth algebraic subvariety of $\mathbf{CP}^n$ is also a symplectic manifold.

3.7. The last sections are a very brief introduction to calculus of variations and its relation to symplectic geometry.

To fix ideas, consider parametric curves $\gamma(t),\ a \le t \le b$ in a manifold M. A *functional* $F : \gamma \to \mathbf{R}$ is a function whose argument is a curve.

Example. Let M be a submanifold in Euclidean space. The length functional is

$$\int_a^b |d\gamma/dt|dt,$$

and the energy functional is

$$\int_a^b |d\gamma/dt|^2 dt.$$

Exercise. Show that the length functional, unlike the energy one, is independent of parameterization of the curve.

An infinitesimal deformation (or a *variation*) of a curve $\gamma(t)$ is a 1-parameter family of curves $\gamma_\varepsilon(t)$; a variation is determined by the vector field along the curve

$$v(t) = d\gamma_\varepsilon(t)/d\varepsilon\ |_{\varepsilon=0}.$$

Informally, $\gamma + \varepsilon v$ is thought of as a curve, close to γ.

Definition. The functional is called differentiable if the familiar limit

$$\Phi(\gamma, v) := (F(\gamma + \varepsilon v) - F(\gamma))/\varepsilon\ |_{\varepsilon=0}$$

is a linear function of v. This linear function of v is often called the variation of F. A curve $\gamma(t)$ is called an *extremum* of F if $\Phi(\gamma, v) = 0$ for every variation v.

Example. A straight segment with fixed end-points in Euclidean space is an extremum of the length functional. Indeed, $\gamma(t) = et$ where e is unit a unit vector. Let $v(t)$ be a variation; note that v vanishes at the end-points of the curve. Then

$$|\gamma_\varepsilon'(t)| = 1 + \varepsilon\ e \cdot v'(t) + O(\varepsilon^2).$$

Since

$$\int e \cdot v'(t)\ dt = 0,$$

the result follows. The extremals of the length functional are called *geodesics*.

Exercise. Prove that the great circles on a sphere are geodesics.

Let $L(x, u, t)$ be a time-dependent smooth function on TM called, in this context, the *Lagrangian*. Consider the functional on, say, closed curves in M:

$$F(\gamma) = \int L(\gamma(t), \gamma'(t), t)\ dt.$$

Theorem. *A curve $\gamma(t)$ is an extremal of the functional F if and only if the following Euler-Lagrange equation holds identically along γ (that is, $x = \gamma(t), u = \gamma'(t)$):*

$$\frac{d}{dt}\frac{\partial L}{\partial u} = \frac{\partial L}{\partial x}.$$

Proof. Choose a local coordinate system near γ. One has:

$$\int L(\gamma + \varepsilon v, \gamma' + \varepsilon v', t)\ dt = \int L(\gamma, \gamma', t)\ dt + \varepsilon \int \Big(\frac{\partial L}{\partial x}(\gamma, \gamma', t)v + \frac{\partial L}{\partial u}(\gamma, \gamma', t)v'\Big)\ dt.$$

Integrating by parts on the RHS, one concludes:

$$\Phi(\gamma, v) = \int \Big(\frac{\partial L}{\partial x}(\gamma, \gamma', t) - \frac{d}{dt}\ \frac{\partial L}{\partial u}(\gamma, \gamma', t)\Big)v\ dt.$$

This integral vanishes for every v iff the expression in the parenthesis is zero, i.e., if Euler-Lagrange equation holds.

Example. Let $L(x, u, t) = u^2/2$. Then the Euler-Lagrange equation reads: $u' = 0$. Therefore the extremals are the lines $x(t) = ut + c$.

Remarks. 1. Assume that $L(x, u)$ does not depend on time. Using the chain rule, the Euler-Lagrange equation can be rewritten as $L_{ux}u + L_{uu}u' = L_x$.

2. Although the computation was made in local coordinates, its result, the Euler-Lagrange equation, has an invariant meaning.

Exercises. 1. Consider a Lagrangian $L(x, u)$ in Euclidean space satisfying:
(i) $L(x, u)$ is homogeneous of degree 1 in u, i.e., $L(x, tu) = tL(x, u)$ for all $t > 0$;
(ii) the matrix of mixed partial derivatives L_{xu} is symmetric.
Prove that the extremals are straight lines (not necessarily arc-length parameterized!)

2. Consider the Lagrangian

$$L(x, u) = \frac{u^2}{1 - x^2} + \frac{(u \cdot x)^2}{(1 - x^2)^2}$$

where $|x| < 1$. Prove that the extremals are straight lines (this is the hyperbolic metric inside the ball).

3. Let M be a hypersurface in Euclidean space. Prove that the geodesics $\gamma(t)$ on M satisfy the following condition: for every t the acceleration vector $\gamma''(t)$ belongs to the 2-plane spanned by the velocity $\gamma'(t)$ and the normal vector to M at point $\gamma(t)$. Also prove that the extremals of the energy functional are the geodesics with a constant speed parameterization. This means that the geodesics are the trajectories of a free particle on M: the only force acting on the particle is perpendicular to the hypersurface.

3.8. Assume that the Lagrangian $L(x, u)$ is a time-independent convex function of u. Again working in local coordinates, set:

$$q = x, p = L_u.$$

In these new coordinates, the Euler-Lagrange equation reads: $p' = L_q$. Define a new (Hamiltonian) function as follows:

$$H(q, p) = pu - L(x, u);$$

the correspondence $L \to H$ is called the *Legendre transformation*; it has an invariant meaning to be discussed in the next section.

Theorem. *The Euler-Lagrange equation is equivalent to the Hamilton equations of the vector field sgrad H:*

$$q' = H_p, p' = -H_q.$$

Proof. One has:

$$dH = H_p dp + H_q dq = pdu + udp - L_x dq - L_u du = udp - L_x dq$$

(the second equality due to $p = L_u$). Therefore

$$H_p = u, \ H_q = -L_x.$$

Since $u = p'$ and, due to the Euler-Lagrange equation, $L_x = d(L_u)/dt = p'$, the result follows.

Remark. The respective symplectic structure is $dp \wedge dq = d(L_u) \wedge dx$.

Thus the trajectories of the Hamiltonian vector field are the extremals of the respective Lagrangian. This result provides an equivalence between Lagrangian and Hamiltonian mechanics, and shows that the variational calculus is closely related to symplectic geometry.

Example: Billiard. Let γ be a closed convex plane curve. The billiard dynamical system describes the motion of a free particle inside γ, subject to the law of elastic reflection: the angle of insidence equals that of reflection. The billiard transformation T acts on the space of oriented lines intersecting γ; it sends an incoming ray to the outgoing one.

Let t be the arc-length parameter along γ. Given an oriented line l, let $\gamma(t_1)$ and $\gamma(t_2)$ be the intersection points with the curve, and let $\phi_{1,2}$ be the angles between γ and l at points $\gamma(t_{1,2})$. Note that (t_1, t_2) can be used as coordinates of the line l. Denote by $H(t_1, t_2)$ the distance between points $\gamma(t_1)$ and $\gamma(t_2)$. It follows from the Lagrange multipliers method that

$$\frac{\partial H(t_1, t_2)}{\partial t_1} = -\cos\phi_1, \quad \frac{\partial H(t_1, t_2)}{\partial t_2} = \cos\phi_2.$$

Hence

$$dH = \cos\phi_2 dt_2 - \cos\phi_1 dt_1,$$

and therefore

$$\sin\phi_2 dt_2 \wedge d\phi_2 = \sin\phi_1 dt_1 \wedge d\phi_1.$$

This is a T-invariant symplectic form.

Exercise. Prove that this T-invariant symplectic form coincides with the previously discussed area form $dp \wedge d\alpha$ on the space of oriented lines.

Continue with billiards. Consider three consecutive points:

$$(t_1, \phi_1) = T(t_0, \phi_0), \quad (t_2, \phi_2) = T(t_1, \phi_1).$$

It folows from the previous formulas that

$$\frac{\partial H(t_0, t_1)}{\partial t_1} + \frac{\partial H(t_1, t_2)}{\partial t_1} = 0.$$

This formula gives the billiard transformation a variational meaning. Suppose one wants to shoot the billiard ball from a given point $x \in \gamma$ so that after a reflection at point $y \in \gamma$ it arrives at a given point $z \in \gamma$. How does one find the unknown point y? Answer: this point is a critical point of the function $|xy| + |yz|$.

Exercise. Extend the above results to multi-dimensional billiards.

3.9. In this last section we discuss the geometrical meaning of the Legendre transformation. For simplicity, we talk about single variable although everything extends verbatum to multivariable case. Let $f(u)$ be a convex function: $f''(x) > 0$. Fix a number p and consider the line $v = pu$. Consider the distance (along the vertical) from the line to the graph of the function:

$$pu - f(u) := F(p, u).$$

There exists a unique u for which this distance is maximal: $f'(u) = p$. Denote this maximal distance by $g(p)$ – this is the Legendre transformation of the function $f(u)$.

Example. If $f(u) = u^a/a$ then $g(p) = p^b/b$ where $1/a + 1/b = 1$.

It follows from the definition that

$$pu \leq f(u) + g(p)$$

for every pair (u, p). This is called the *Young inequality*. In particular,

$$pu \leq \frac{u^a}{a} + \frac{p^b}{b}$$

for $u, p > 0,\ a, b > 1,\ 1/a + 1/b = 1$.

Proposition. *The Legendre transformation is involutive.*

Proof. Consider two planes: (u, v) and (p, r) ones. Consider the equation

$$r + v = pu.$$

For a fixed (p, r) this equation describes a non-vertical line in the first plane; for a fixed (u, v) it describes a non-vertical line in the second one. Thus the planes are *dual*: a (non-vertical) line in one is a point of the other.

Given a convex curve γ in one plane, the set of its tangent lines is a curve γ^* in the other plane; this curve is called a *dual curve*. It follows from definition of the Legendre transformation that the line $v = pu - g(p)$ is tangent to the graph $v = f(u)$. Thus the graph of $g(p)$ is dual to the graph of $f(u)$.

It remains to show that duality is an involutive relation. Let γ be a curve and γ^* its dual. A tangent line to γ^* is a limit position of a line l through two very close points, say, $A, B \in \gamma^*$. The points A, B in the second plane correspond to two lines a, b in the first plane, tangent to γ. The intersection of two very close tangents to γ is a point L, dual to the line l and close to γ; in the limit, $L \in \gamma$. Thus $\gamma^{**} = \gamma$.

Exercise. Describe the curves, dual to $v = u^2$ and to $v = u^3$.