Introduction to differentiable manifolds and symplectic geometry

Tilmann Wurzbacher

Institut de Recherche Mathématique Avancée

Université Louis Pasteur et C.N.R.S.

7, rue René Descartes

F-67084 Strasbourg Cedex, France

email: wurzbach@math.u-strasbg.fr

web: http://www-irma.u-strasbg.fr/~wurzbach/

Abstract:

Assuming only undergraduate level knowledge of linear algebra, analysis including ordinary differential equations and rudimentary topology, we develop the basics of the theory of symplectic manifolds and Hamiltonian dynamical systems, that is pivotal in geometric considerations in theoretical physics. The material on multilinear and symplectic algebra as well as on differentiable manifolds (vector fields, differential forms and de Rham cohomology) necessary to bridge the gap from the prerequisites to symplectic geometry is thoroughly covered in the first chapters of the text.

Table of contents

Introduction

These lecture notes are based on courses I gave in Villa de Leyva in Colombia and in Hamburg in Germany, mainly for students of mathematics and/or theoretical physics, during the second half of the year 1999. Since most of the more advanced material of the course in Villa de Leyva is available as part of the text [Wu], I concentrated here on the foundations of the theory of differentiable manifolds (and of symplectic geometry), being at the base of most considerations in the field of geometry related to theoretical physics.

After a short motivation of the Hamiltonian approach to mechanics, the main body of the text proceeds as follows:

In Chapter 1, we complement standard knowledge in linear algebra by a thorough development of multilinear algebra, indispensable for the calculus of differential forms on manifolds, as well as of "symplectic algebra", i.e., the basic results on symplectic vector spaces as, e.g., normal forms and the existence of compatible complex structures.

The second chapter develops the theory of finite dimensional manifolds from scratch. We give complete proofs of all crucial points of the text, with the only exceptions of the construction of partitions of unity, the proof of Stokes' theorem and of "Moser's formula". We include de Rham cohomology in our presentation since it plays a prominent role in theoretical physics (and of course in geometry), though admittedly physicists tend to describe it in a different language. We strongly believe that learning the general mathematical formulation at an early stage is well worth the effort since it unifies several important notions.

Chapter 3 is an introduction to symplectic geometry and Hamiltonian dynamical systems. The concise formulation and easy proofs of the foundational results of analytical mechanics as, e.g., the theorems of Darboux and Noether, the existence of symplectic structures on the total space of the cotangent bundle of a manifold and the properties of the Poisson structure on a symplectic manifolds here show the usefulness of the preceding chapter. We also give some basic material of contemporary symplectic differential geometry as the notion of a Kählerian manifold and rudiments of the theory of non-linear symplectic maps.

Though there are a few references scattered throughout the text, we conclude each section with some "Bibliographical remarks", where hints on related literature are given, as an incitation for further reading and self-study.

Let me take the opportunity to thank all participants of the courses in Villa de Leyva and Hamburg for their interest and "feedback". Last but not least I would like to thank Sergio Adarve, Dorothea Glasenapp, Sylvie Paycha, Peter Slodowy, Andrés Reyes, Rolando Roldán and Mónica Vargas without whose efforts these courses and lecture notes would not have been possible and with whom working together was always a pleasure for me.

Tilmann Wurzbacher

Strasbourg, June 26, 2000.

0. Motivation

0.1 An example from mechanics: from Newton to Lagrange to Hamilton

Let us consider a particle of mass m ($m > 0$, small compared to the mass of the earth), "close" to the earth, subject to the gravitational field of the earth. We can assume the surface of the earth to be the plane $q_3 = 0$ and describe the trajectory of the particle by $q(t) = (q_1(t), q_2(t), q_3(t))$ with $q_3 \geq 0$. The exterior force acting on the particle is given by $F = -mg\ e_3$ with g denoting the (strictly positive) "gravitational constant" and e_3 standing for the third unit vector in $\mathbb{R}^3$.

Newtonian description:

The Newtonian equation of motion is

$$F = m\ddot{q}$$

with initial conditions

$$q(0) = q^0 \text{ with } q_3^0 \geq 0 \text{ and } \dot{q}(0) = v^0 .$$

Here we have $-mge_3 = m\ddot{q}$ so that the trajectory is given by

$$q(t) = q^0 + tv^0 - g\frac{t^2}{2}e_3 .$$

Observation. The above force field $\vec{F}$ is "conservative", i.e. there is a function

$$U : \{q \in \mathbb{R}^3 \,|\, q_3 \geq 0\} \to \mathbb{R} \text{ such that } \vec{F} = -\vec{\nabla} U .$$

This function, here we can take e.g. $U(q) = mgq_3$, is called the "potential energy", whereas the function $T = \frac{m}{2}(\dot{q})^2 = \frac{m}{2}((\dot{q}_1)^2 + (\dot{q}_2)^2 + (\dot{q}_3)^2)$ is considered as the "kinetic energy". In the case of a conservative force field we can go to the

Lagrangian description:

Let $L = L(q, \dot{q}, t) := T - U$ be the "Lagrange function" (that might depend explicitly on t!), then the Lagrangian equation of motion is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0 \text{ for } k = 1, 2, 3$$

with initial condition as above

$$q(0) = q^0 \text{ with } q_3^0 \geq 0 \text{ and } \dot{q}(0) = v^0 .$$

In our example these equations are obviously the same as in the Newtonian approach.

Observation. We have – at least in our example –

$$\frac{\partial L}{\partial \dot{q}_k} = m\dot{q}_k = p_k ,$$

the "(linear) momentum" of the particle. We can thus write the "total energy" $H = H(q, p) = T + U$ as a function of q and p! (This transition from L to H is called "Legendre transformation" and is **not** always possible! Since later on we will not be concerned with Lagrangian mechanics we do not go into this question more deeply.)

For such a "Hamilton function" or "Hamiltonian" H we have the following Hamiltonian equations of motion

$$\frac{\partial H}{\partial p_k} = \dot{q}_k \quad \text{and} \quad \frac{\partial H}{\partial q_k} = -\dot{p}_k \text{ for } k = 1, 2, 3$$

with initial condition

$$q(0) = q^0 \text{ and } p(0) = p^0 .$$

In the example we have $H = \frac{1}{2m}p^2 + mgq_3$ (setting of course $p^2 = (p_1)^2 + (p_2)^2 + (p_3)^2$) and $p(0) = mv^0$. The equation of motion then reads as follow

$$\dot{q}_k = \frac{\partial H}{\partial p_k} = \frac{1}{m}p_k \quad \text{and} \quad \dot{p}_k = -\frac{\partial H}{\partial q_k} = -mg\delta_{k,3} .$$

Differentiating the first equation with respect to time t and inserting the result into the second we find the Newtonian equation of motion for $q = q(t)$ and we get $p = p(t)$ then in the case of the example trivially from $q(t)$ and the first equation.

Remark. The choice between the Lagrangian and the Hamiltonian approach (in a situation where both can be applied) depends on further details: the advantages of the Hamiltonian approach lie in the equal treatment of the variables q and p, the first order of the equations, and a simple transition to quantum mechanics. On the other hand, in relativistic mechanics or in the transition from classical field theory to quantum field theory the Lagrangian might often be more useful, at least for theoretical physicists.

0.2 An infinite dimensional example: the wave equation

Without going into the (functional-analytic) problems of domains crucial in infinite dimensional situations we will give here a simple class of a Hamiltonian equations, which are "equivalent" to certain non-linear wave equations.

Let for $n \geq 1$

$$\mathcal{S}(\mathbb{R}^n) = \left\{ f : \mathbb{R}^n \to \mathbb{R} \mid \lim_{||x|| \to \infty} P(x) Q\left(\frac{\partial}{\partial x}\right) f(x) = 0 \; \forall P, Q \in \mathbb{R}[T_1, ..., T_n] \right\}$$

(with $Q(\frac{\partial}{\partial x}) = Q(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n})$ viewed as a scalar partial differential operator with constant coefficients). Given a function U in, e.g., $\mathcal{S}(\mathbb{R})$ we define the vector space $E = \mathcal{S}(\mathbb{R}^3) \oplus \mathcal{S}(\mathbb{R}^3)$ and a function $H : E \to \mathbb{R}$ by

$$H(\phi, \pi) = \int_{\mathbb{R}^3} \left(\frac{\pi^2}{2} + \frac{\|\nabla \phi\|^2}{2} + U(\phi) \right) d^3 x \, .$$

By analogy we can consider the following Hamiltonian equation

$$\frac{\partial H}{\partial \pi} = \dot{\phi} \quad \text{and} \quad \frac{\partial H}{\partial \phi} = -\dot{\pi} \, .$$

We interpret the partial derivatives as "L^2-gradients" in the following sense

$$\left\langle \left(\frac{\partial H}{\partial \pi} \right) (\phi_0, \pi_0), \tilde{\pi} \right\rangle_{L^2(\mathbb{R}^3)} = (D_2 H)_{(\phi_0, \pi_0)} (\tilde{\pi}) = \frac{d}{d\epsilon}\Big|_0 H(\phi_0, \pi_0 + \epsilon \tilde{\pi})$$

and analogously for ϕ.

A direct calculation yields:

$$\left\langle \left(\frac{\partial H}{\partial \pi} \right) (\phi_0, \pi_0), \tilde{\pi} \right\rangle_{L^2(\mathbb{R}^3)} = \langle \pi_0, \tilde{\pi} \rangle_{L^2(\mathbb{R}^3)} \quad \text{and}$$

$$\left\langle \left(\frac{\partial H}{\partial \phi} \right) (\phi_0, \pi_0), \tilde{\phi} \right\rangle_{L^2(\mathbb{R}^3)} = \left\langle -\Delta \phi_0 + U'(\phi_0), \tilde{\phi} \right\rangle_{L^2(\mathbb{R}^3)} ,$$

where U' denotes the derivative of U.

In the case at hand the Hamiltonian equation thus reads as follows:

$$\dot{\phi} = \frac{\partial H}{\partial \pi} = \pi \quad \text{and} \quad \dot{\pi} = -\frac{\partial H}{\partial \phi} = \Delta \phi_0 - U'(\phi_0) \, .$$

(We suppress here the discussion of the initial condition as well as the existence and uniqueness questions for the solutions of this equation.)

Thus a curve $t \mapsto (\phi, \dot{\phi}) \in E$ satisfies the Hamilton equation if and only if the function $(t, x) \mapsto \phi(t)(x)$ fulfills the following nonlinear wave equation on $\mathbb{R}^3$:

$$\ddot{\phi} = \Delta\phi - U'(\phi)\,.$$

This example indicates that techniques from Hamiltonian systems might be useful in the study of partial differential equations. Other applications of infinite dimensional symplectic geometry and Hamiltonian systems might come from the extension of geometric quantization (or any other quantization procedure starting from Hamiltonian mechanics) to infinite dimensional situations.

0.3 Solving the Hamilton equation for one degree of freedom

Example. In the theory of small oscillations the basic Newtonian equation is as follows:

$$\ddot{q} = F(q) = -m\omega^2 q \text{ with } m > 0 \text{ and } \omega > 0\,.$$

The corresponding Hamilton function is

$$H : \mathbb{R} \to \mathbb{R},\ H(q,p) = \frac{1}{2m}p^2 + \frac{m\omega^2}{2}q^2\,,$$

and the Hamiltonian equation is given by

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{1}{m}p \text{ and } \dot{p} = -\frac{\partial H}{\partial q} = -m\omega^2 q$$

with initial condition

$$q(0) = q^0 \text{ and } p(0) = p^0\,.$$

This equation is equivalent to the Newtonian equation $\ddot{q} = -\omega^2 q$ and its solution is easily calculated:

$$q(t) = q^0\cos(\omega t) + \frac{p^0}{m\omega}\sin(\omega t),\ p(t) = p^0\cos(\omega t) - q^0 m\omega\sin(\omega t)\,.$$

Remarks. Let V be open in $\mathbb{R}^{2n}$ and $H = H(q,p)$ a real-valued smooth function on V.

(1) "Conservation of energy"

For each (local) solution $t \mapsto \gamma(t) = (q(t), p(t))$ of the Hamiltonian equation one has:

$$\frac{d}{dt}H(\gamma(t)) = 0\,,$$

i.e. for a connected open intervall I and any t_0 in I one has

$$H(\gamma(t)) = H(\gamma(t_0)) = H^0 \text{ for all } t \in I\,.$$

(Proof as an exercise.)

(2) A constant solution of the Hamilton equation is called a “stationary solution” or an “equilibrium point”.

(3) For a Hamilton function of the type

$$H(q,p) = T + U(q) = \frac{1}{2m}p^2 + U(q)$$

(with $p^2 = \|p\|^2$ as usual) a (local) solution of the Hamilton equation with initial condition $q(t_0) = q^0$ and $p(t_0) = p^0$ is constant if and only if the following holds:

$$\big(p^0 = 0 \ \text{ and } \ (\nabla U)(q^0) = 0\big) \ .$$

(Proof again as an exercise.)

Solution of the Hamilton equation: one degree of freedom.

Let us consider here only Hamiltonians of the type $H = \frac{1}{2m}p^2 + U(q)$ with q, p in an open set in $\mathbb{R}^2$. (We denote here the dimension of the “configuration space”, i.e. “the space of q's”, as the number of “degrees of freedom” of a mechanical system; see below in Section 3.1.)

Case 1. Equilibrium points

From Remark (3) above we know that equilibrium points are caracterized by $(p^0 = 0$ and $(\nabla U)(q^0) = 0)$. Obviously the solutions are then given by $\gamma(t) \equiv \gamma(t_0) = (q^0, p^0)$ for all t in $\mathbb{R}$.

Case 2. $p^0 \neq 0$

Since in this case $T(p^0) > 0$ we know by the conservation of energy that $E_0 - U(q(t)) > 0$ for t near t_0 on a solution. (We denote $H(q^0, p^0)$ by E_0 here.)

It follows that $p(t) = \pm\sqrt{2m(E_0 - U(q(t))}$ and for $t - t_0$ small the sign is given by the sign of p^0.

Let us assume without loss of generality that $p^0 > 0$. The Hamilton equation implies now

$$\frac{dq}{dt} = \sqrt{\frac{2}{m}(E_0 - U(q(t))}\,.$$

Since the right hand side is nonvanishing we can write

$$dt = \frac{dq(t)}{\sqrt{\frac{2}{m}(E_0 - U(q(t)))}}$$

and we thus arrive, at least for $t_1 - t_0$ small, at the following equation:

$$t_1 - t_0 = \int_{t_0}^{t_1} dt = \int_{t_0}^{t_1} \frac{dq(t)}{\sqrt{\frac{2}{m}(E_0 - U(q(t)))}}\,.$$

For t in $[t_0, t_1]$ the function $t \mapsto q(t)$ is invertible and hence

$$t_1 - t_0 = \int_{q(t_0)}^{q(t_1)} \left(\frac{1}{\sqrt{\frac{2}{m}(E_0 - U(q))}} \right) dq = \int_{q(t_0)}^{q(t_1)} f(q) dq = F(q(t_1)) - F(q(t_0)),$$

where $F' = f$. This is equivalent to

$$F(q(t)) = t + (F(q(t_0)) - t_0) \quad \text{for} \quad |t - t_0| \quad \text{small.}$$

Furthermore we observe that $F'(q^0) = \frac{m}{p^0} \neq 0$ and thus F is locally invertible near $q(t_0) = q^0$, i.e., $F^{-1} = G$. This allows the following formula, obtained only by "quadratures" (i.e. integration of functions in one variable) and "algebraic operations" (i.e. the usual operations on numbers and the calculation of inverses of functions in one variable):

$$q(t) = G(t + F(q^0) - t_0), \; p(t) = m\dot{q}(t)$$

for t sufficiently close to t_0.

Case 3. $p^0 = 0$ and $\frac{\partial U}{\partial q}(q^0) \neq 0$

Since we are not allowed to divide by $\sqrt{2/m(E_0 - U(q))}$ as in the preceding case, we have to use the condition $\frac{\partial U}{\partial q}(q^0) \neq 0$ that implies that there exists a function $Q = Q(u)$ locally defined near $u^0 = U(q^0)$ such that $Q \circ U(q) = q$.

Using conservation of energy we have

$$U(q(t)) = E_0 - \frac{p(t)^2}{2m} =: f_1(p(t))$$

on solutions and thus we have (near p^0)

$$q(t) = (Q \circ f_1)(p(t))$$

for solutions of the Hamilton equation.

This equation implies furthermore

$$\dot{p} = -\frac{\partial U}{\partial q} = -U'(q) = ((-U') \circ Q \circ f_1)(p) =: f_2(p)$$

and $f_2(p^0) = (-U')(q^0) \neq 0$ and thus we have, for p close to p^0 that $f_2(p) \neq 0$.

Using $f(p) = \frac{1}{f_2(p)}$ we rewrite this for p near p^0 as follows:

$$dt = \frac{dp(t)}{f_2(p(t))} = f(p(t))dp(t).$$

We deduce for $|t_1 - t_0|$ sufficiently small that

$$t_1 - t_0 = \int_{t_0}^{t_1} dt = \int_{t_0}^{t_1} f(p(t))dp(t) = \int_{p(t_0)}^{p(t_1)} f(p)dp,$$

since $\dot{p}(t_0) = \frac{\partial U}{\partial q}(q^0) \neq 0$ implies that $t \mapsto p(t)$ is a variable transformation for t close to t_0. Thus, with F a primitive of f, we have $t_1 - t_0 = F(p(t_1)) - F(p(t_1))$ and therefore

$$F(p(t)) = t + F(p^0) - t_0$$

for t close to t_0. Since $F'(p) = f(p) = \frac{1}{f_2(p)} \neq 0$ for p close to p^0 and thus – a fortiori – for $p(t)$ with t close to t_0, there exists a local inverse $G = F^{-1}$ with $G \circ F(p) = p$ for p close to p^0.

We arrive at the "explicit" solution formula

$$p(t) = G(t + F(p^0) - t_0)$$

and with $\dot{q} = \frac{p}{m}$ we reach the following conclusion

$$q(t) = \int_{t_0}^{t_1} \frac{p(s)}{m} ds + q^0 = \int_{t_0}^{t_1} \frac{G(s + F(p^0) - t_0)}{m} ds + q^0 .$$

Back to the example of the harmonic oscillator.

Since $\frac{\partial U}{\partial q} = m\omega^2 q$ we find that

$$\frac{\partial U}{\partial q} = 0 \text{ if and only if } q = 0 .$$

Case 1. Equilibrium points

The only equilibrium point is given by $p^0 = 0, q^0 = 0$.

Case 2. $p^0 \neq 0$

Let us without loss of generality assume that $p^0 > 0$.

The condition $E_0 - U(q) > 0$ is equivalent to $|q| < \sqrt{\frac{2E_0}{m\omega^2}}$ and for these q we have

$$f(q) = \sqrt{\frac{m}{2E_0}} \frac{1}{\sqrt{1 - \left(\frac{m\omega^2}{2E_0}\right) q^2}}$$

with a primitive given by

$$F(q) = \frac{1}{\omega} \arcsin\left(\sqrt{\frac{m\omega^2}{2E_0}} \cdot q\right) .$$

For $|r| < \frac{\pi}{2\omega}$ we have the following explicit inverse of F:

$$G(r) = \sqrt{\frac{2E_0}{m\omega^2}} \sin(\omega r) .$$

Thus the local solution of the Hamilton equation is

$$q(t) = G(t + F(q^0)) = \sqrt{\frac{2E_0}{m\omega^2}} \sin(\omega t + \omega F(q^0))\,.$$

A direct calculation shows that

$$q(t) = q^0 \cos(\omega t) + \frac{p^0}{m\omega} \sin(\omega t)$$

for $|q| < \sqrt{\frac{2E_0}{m\omega^2}}$, i.e. as long as $p(t) > 0$.

Observation. The above solution is a priori only locally defined, i.e. for t close to t_0. In the case at hand we can immediately extend it to a solution for all real t.

Case 3. $p^0 = 0$ and $\frac{\partial U}{\partial q}(q^0) \neq 0$

We leave the analysis of this case as an exercise.

Remark. The aim to find explicit formulas for the solutions of Hamiltonian systems led to the discovery of many important special functions in the 19th century. Notably the theory of analytic functions in one complex variable and of "Riemann surfaces" was highly stimulated by this search.

Bibliographical remarks. *Our - highly subjective - choice of physics texts on classical mechanics include [Ar1], [Gol] and [Sch]. For the mathematical approach to infinite dimensional Hamiltonian mechanics see, e.g., [AMR] and [CM]. A good german reference is [Lau] which we followed in Section 0.2.*

1. Multilinear and symplectic algebra

In Chapter 1 all vector spaces will be finite dimensional over a field $\mathbb{K}$ which is $\mathbb{R}$ or $\mathbb{C}$ if not explicitely stated otherwise.

1.1 Multilinear forms

Definition. A "bilinear form" on a $\mathbb{K}$–vector space V is a map $B : V \times V \to \mathbb{K}$ such that

(i) $B(v + v', w) = B(v, w) + B(v', w)$

$B(\lambda \cdot v, w) = \lambda \cdot B(v, w)$ and

(ii) $B(v, w + w') = B(v, w) + B(v, w')$

$B(v, \mu \cdot w) = \mu \cdot B(v, w)$ for all v, v', w, w' in V and for all λ, μ in $\mathbb{K}$.

Remark. Using the canonical basis $\{e_1, ..., e_n\}$ of $\mathbb{K}^n$ a bilinear form B on $\mathbb{K}^n$ can be represented in a unique way by a square matrix $Q = Q_B$ in $\mathrm{Mat}(n \times n, \mathbb{K})$ as follows:

$$B(x,y) = B\Big(\sum_{j=1}^{n} x_j e_j, \sum_{k=1}^{n} y_k e_k\Big) = \sum_{j,k=1}^{n} x_j B(e_j, e_k) y_k = \sum_{j,k=1}^{n} x_j Q_{jk} y_k = {}^t x \cdot Q_B \cdot y.$$

Lemma. *The map* $\mathcal{B}(V) := \{B : V \times V \to \mathbb{K} \,|\, B \text{ is bilinear}\} \to \ \mathrm{Mat}(n \times n, \mathbb{K}), B \mapsto Q_B$ *is a* $\mathbb{K}$*–vector space isomorphism.*

Proof. Exercise. □

Remark. Let $T : U \to V$ be a $\mathbb{K}$–linear map and B a bilinear form on V, then we define the "pullback of B unter T" by

$$(T^*B)(u_1, u_2) := B(T(u_1), T(u_2))$$

for all u_1, u_2 in U. We observe that T^*B is a bilinear form von U.

Special Case. Let $U = V = \mathbb{K}^n$ and $T = T_A$, the linear map $x \mapsto A{\cdot}x$ associated to a $(n \times n)$–matrix A. Then $Q_{T^*B} = {}^t A \cdot Q_B \cdot A$.

Proof. Exercise. □

Definition. (1) A bilinear form B on V is called "symmetric"if

$$B(v, v') = B(v', v) \quad \text{for all } v, v' \text{ in } V.$$

(2) A bilinear form B on V is called "skew–symmetric" (or "anti-symmetric" or "alternating") if

$$B(v, v') = -B(v', v) \quad \text{for all } v, v' \text{ in } V.$$

Remark. If B is skew–symmetric, then $B(v, v) = 0$ for all v in V.

Lemma. *(i) Each bilinear form B on V is uniquely decomposed into the sum of a symmetric and an alternating bilinear form.*

(ii) *If B is a bilinear form on $\mathbb{K}^n$ and Q_B the associated matrix, then one has*

$$(B \textit{ is symmetric if and only if } \ {}^t(Q_B) = Q_B)$$

and

$$(B \textit{ is skew–symmetric if and only if } \ {}^t(Q_B) = -Q_B).$$

Proof. Exercise. □

Let us recall the "tensor algebra" language: we denote by $\bigotimes^k V^*$ the set of multilinear maps from $V^k = \underbrace{V \times \cdots \times V}_{k \text{ times}} \to \mathbb{K}$, and by $\mathcal{T}(V^*)$ the "tensor algebra (over V^*)":

$$\bigoplus_{k \geq 0} \bigotimes^k V^* = \{(m_0, m_1, \ldots) | m_k \in \ \otimes^k V^* \text{ and } m_k = 0 \text{ for almost all } \ k\}.$$

If $\{\epsilon_1, \ldots, \epsilon_n\}$ is an ordered basis of V and $\{\epsilon_1^*, \ldots, \epsilon_n^*\}$ the dual basis of V^* then

$$\{\epsilon_{i_1}^* \otimes \ldots \otimes \epsilon_{i_k}^* |\, i_1, \ldots, i_k \in \{1, \ldots, n\}\}$$

is a basis of $\bigotimes^k V^*$ and thus its dimension equals n^k. The multiplication on $\mathcal{T}(V^*)$ is given as follows: let $t \in \otimes^k V^*, s \in \otimes^l V^*$ and $v_1, \ldots, v_{k+l} \in V$. Then the element $t \bigotimes s$ of $\bigotimes^{k+l} V^*$ is defined by

$$(t \otimes s)(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+l}) := t(v_1, \ldots, v_k)s(v_{k+1}, \ldots, v_{k+l}).$$

Lemma. *The $\mathbb{K}$–vector space $\mathcal{T}(V^*)$ together with the multiplication given by $\bigotimes$ is a non–commutative, associative, unital $\mathbb{K}$–algebra.*

Proof. Exercise. $\square$

The “symmetric group” S_k of all permutations of the set $\{1, \ldots, k\}$ acts on $\bigotimes^k V^*$ as follows:

$$\sigma(t)(v_1, \ldots, v_k) := t(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \quad \forall v_1, \ldots, v_k \in V.$$

Lemma. *Let σ and τ be in S_k and t in $\bigotimes^k V^*$. Then*

$$(\sigma \circ \tau)(t) = \sigma(\tau(t)).$$

Proof. Setting $w_k := v_{\sigma(k)}$ we have

$$\sigma(\tau(t))(v_1, \ldots, v_k) = \tau(t)(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = \tau(t)(w_1, \ldots, w_k) = t(w_{\tau(1)}, \ldots, w_{\tau(k)}).$$

By definition we have $v_{\sigma(\tau(j))} = w_{\tau(j)}$ and thus

$$t(w_{\tau(1)}, \ldots, w_{\tau(k)}) = t(v_{\sigma(\tau(1))} \ldots, v_{\sigma(\tau(k))}) = t(v_{(\sigma\circ\tau)(1)}, \ldots, v_{\sigma\circ\tau(k)}) = (\sigma \circ \tau)(t)(v_1, \ldots, v_k).$$

$\square$

Lemma. *Let $\chi : S_k \to \mathbb{K}\backslash\{0\}$ be a homomorphism. Then either $\chi(\sigma) = 1$ for all σ in S_k or $\chi(\sigma) = \mathrm{sign}(\sigma)$, defined by $(-1)^r$ on a product $\sigma = \sigma_1 \circ \cdots \circ \sigma_r$ of transpositions σ_j for $j = 1, \ldots, r$.*

Idea of the proof. The group S_k is generated by transpositions and $\mathbb{K}\backslash\{0\}$ is abelian. $\square$

Definition. (1) The space of “symmetric k–forms” is given by

$$\mathcal{S}^k V^* = \{t \in \otimes^k V^* | \sigma(t) = t \quad \forall \sigma \in S_k\}.$$

(2) The space of “skew–symmetric (or alternating) k–forms” is given by

$$\Lambda^k V^* = \{t \in \otimes^k V^* | \sigma(t) = \mathrm{sign}(\sigma) \cdot t \quad \forall \sigma \in S_k\}.$$

Remarks. (1) We have a natural “symmetrizer” map

$$\mathrm{Symm} : \otimes^k V^* \to \otimes^k V^*,\ \mathrm{Symm}(t) = \frac{1}{k!} \sum_{\sigma \in S_k} \sigma(t)$$

such that Symm $\circ$ Symm $=$ Symm and the image of Symm is $\mathcal{S}^kV^*$. Obviously Symm extends to a map $\mathcal{T}(V^*) \to \mathcal{T}(V^*)$ with analogous properties. Combining Symm with the tensor product $\otimes$ we obtain a multiplication on

$$\mathcal{S}(V^*) = \bigoplus_{k\geq 0} \mathcal{S}^kV^* : \ t \vee s := \mathrm{Symm}(t\otimes s) \quad \text{for } t, s \text{ in } \mathcal{S}(V^*).$$

It follows that $\mathcal{S}(V^*)$ is a commutative, associative, unital $\mathbb{K}$–algebra. Let us also remark that the dimension of $\mathcal{S}^kV^*$ equals $\begin{pmatrix} n+k-1 \\ k \end{pmatrix}$ if n is the dimension of V, and that $\mathcal{S}(V^*)$ is isomorphic to the space of polynomials on V as a $\mathbb{K}$–algebra. (See [Gre] for proofs of this and more details on symmetric tensors.)

(2) Analogously, we have a natural "alternator" or "anti–symmetrizer" map

$$\mathrm{Alt} : \otimes^k V^* \to \otimes^k V^*, \quad \mathrm{Alt}(t) = \frac{1}{k!}\sum_{\sigma\in S_k} \mathrm{sign}\ (\sigma)\sigma(t)$$

such that Alt $\circ$ Alt $=$ Alt and $\mathrm{Alt}(\otimes^kV^*) = \Lambda^kV^*$. Again Alt extends to a map

$$\mathrm{Alt} : \mathcal{T}(V^*) \to \Lambda(V^*) = \bigoplus_{k\geq 0}\Lambda^kV^*.$$

Definiton. Let α be in $\otimes^kV^*$ and β in $\otimes^lV^*$, then the "wedge product of α with β" is defined as follows:

$$\alpha\wedge\beta := \frac{(k+l)!}{k!\,l!}\mathrm{Alt}(\alpha\otimes\beta).$$

Remark. The factor in the above definition of the wedge product is chosen such that it relates in the easiest possible way to volumes: let $\{\epsilon_1, \epsilon_2\}$ be a basis of a vector space V and $\{\epsilon_1^*, \epsilon_2^*\}$ the dual basis, then with the above definition

$$\epsilon_1^*\wedge\epsilon_2^*(\epsilon_1, \epsilon_2) = 1.$$

Proposition. *Let α be in $\otimes^kV^*$, β in $\otimes^lV^*$ and γ in $\otimes^mV^*$. Then*

(i) $\alpha\wedge\beta = Alt(\alpha)\wedge\beta = \alpha\wedge Alt(\beta) = Alt(\alpha)\wedge Alt(\beta)$.

(ii) "$\wedge$" is $\mathbb{K}$–bilinear.

(iii) $\alpha\wedge\beta = (-1)^{k\cdot l}\beta\wedge\alpha$.

(iv) $\alpha\wedge(\beta\wedge\gamma) = (\alpha\wedge\beta)\wedge\gamma$.

Proof. *Ad (i).*

$$\mathrm{Alt}(\mathrm{Alt}(\alpha)\otimes\beta) = \mathrm{Alt}\Big(\frac{1}{k!}\sum_{\sigma\in S_k}\mathrm{sign}(\sigma)\sigma(\alpha)\otimes\beta\Big)$$

$$= \frac{1}{(k+l)!}\sum_{\mu\in S_{k+l}}\mathrm{sign}(\mu)\Big(\frac{1}{k!}\sum_{\sigma\in S_k}\mathrm{sign}(\sigma)\mu(\sigma(\alpha)\otimes\beta)\Big).$$

We define σ' in S_{k+l} by setting $\sigma'(j) = \sigma(j)$ for $j = 1, \dots, k$ and $\sigma'(j) = j$ for $j = k+1, \dots, k+l$. It follows

$$\begin{aligned}
\mathrm{Alt}(\mathrm{Alt}(\alpha) \otimes \beta) &= \frac{1}{k!} \sum_{\sigma \in S_k} \left(\frac{1}{(k+l)!} \sum_{\mu \in S_{k+l}} (\mathrm{sign}(\mu \circ \sigma'))(\mu \circ \sigma')(\alpha \otimes \beta) \right) \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{Alt}(\alpha \otimes \beta) = \mathrm{Alt}(\alpha \otimes \beta).
\end{aligned}$$

Analogously one has $\mathrm{Alt}(\alpha \otimes \mathrm{Alt}(\beta)) = \mathrm{Alt}(\alpha \otimes \beta)$ and *(i)* follows easily.

Ad (ii). Since Alt is linear and "$\otimes$" is bilinear, the second assertion follows directly from the definition of "$\wedge$".

Ad (iii). Let $v_1, \dots, v_{l+k}$ be vectors in V.

Then

$$\begin{aligned}
\beta \otimes \alpha(v_1, \dots, v_l, v_{l+1}, \dots, v_{l+k}) &= \alpha \otimes \beta(v_{l+1}, \dots, v_{l+k}, v_1, \dots, v_l) \\
&= \alpha \otimes \beta(v_{\mu_{l,k}(1)}, \dots, v_{\mu_{l,k}(l+k)}),
\end{aligned}$$

where the permutation $\mu_{l,k}$ in S_{l+k} is uniquely defined by the last equality sign and $\mathrm{sign}(\mu_{l,k}) = (-1)^{l \cdot k}$. Thus $\beta \otimes \alpha = \mu_{l,k}(\alpha \otimes \beta)$ and one finds

$$\mathrm{Alt}(\beta \otimes \alpha) = \frac{1}{(k+l)!} \sum_{\tau \in S_{k+l}} \mathrm{sign}(\tau)\tau(\beta \otimes \alpha)$$

$$= \mathrm{sign}(\mu_{l,k}) \frac{1}{(k+l)!} \sum_{\tau \in S_{k+l}} \mathrm{sign}(\tau) \cdot \mathrm{sign}(\mu_{l,k}) \tau(\mu_{l,k}(\alpha \otimes \beta))$$

$$= \mathrm{sign}(\mu_{l,k}) \frac{1}{(k+l)!} \sum_{\tau \in S_{k+l}} (\mathrm{sign}(\tau \circ \mu_{l,k}))(\tau \circ \mu_{l,k})(\alpha \otimes \beta) = (-1)^{k \cdot l} \mathrm{Alt}(\alpha \otimes \beta).$$

Thus $\beta \wedge \alpha = (-1)^{k \cdot l}\ \alpha \wedge \beta$.

Ad (iv). Using (i) we have

$$\begin{aligned}
\alpha \wedge (\beta \wedge \gamma) = \alpha \wedge \left(\frac{(l+m)!}{l!m!} \mathrm{Alt}(\beta \otimes \gamma) \right) &= \frac{(l+m)!}{l!m!} \alpha \wedge (\beta \otimes \gamma) \\
&= \frac{(l+m)!(k+l+m)!}{l!m!k!(l+m)!} \mathrm{Alt}(\alpha \otimes (\beta \otimes \gamma)) \\
= \frac{(k+l+m)!(k+l)!}{m!k!l!} \mathrm{Alt}((\alpha \otimes \beta) \otimes \gamma) &= \frac{(k+l)!}{k!l!} (\alpha \otimes \beta) \wedge \gamma \\
= \left(\frac{(k+l)!}{k!l!} \mathrm{Alt}(\alpha \otimes \beta) \right) \wedge \gamma &= (\alpha \wedge \beta) \wedge \gamma. \qquad \square
\end{aligned}$$

Corollary 1. *If $\{\epsilon_1, \dots, \epsilon_n\}$ is a basis of V with dual basis $\{\epsilon_1^*, \dots, \epsilon_n^*\}$ then*

$$\{\epsilon_{i_1}^* \wedge \ldots \wedge \epsilon_{i_k}^* | 1 \le i_1 < \cdots < i_k \le n\}$$

is a basis of $\Lambda^k V^*$.

Proof. Exercise. $\square$

Corollary 2. *If* $k > n = \dim_{\mathbb{K}} V$ *then* $\Lambda^k V^* = \{0\}$, *and for* $0 \leq k \leq n$ *we have* $\dim_{\mathbb{K}} \Lambda^k V^* = \begin{pmatrix} n \\ k \end{pmatrix}$. *For the whole "exterior algebra" we have*

$$\Lambda(V^*) = \bigoplus_{k \geq 0} \Lambda^k V^* = \bigoplus_{k=0}^{n} \Lambda^k V^*$$

and its dimension equals 2^n.

Proof. Obvious from Corollary 1. $\square$

Let us make a slight digression with the following

Definition. A $\mathbb{K}$–algebra A is called a "$\mathbb{K}$–super algebra" or "$\mathbb{Z}_2$–graded $\mathbb{K}$–algebra" if $A = A_{\bar{0}} \oplus A_{\bar{1}}$ as a $\mathbb{K}$–vector space and the multiplication on A fulfils ($\alpha, \beta \in \{\bar{0}, \bar{1}\} = \mathbb{Z}_2$):

$$a_\alpha \cdot b_\beta \in A_{\alpha+\beta} \quad \forall a_\alpha \in A_\alpha, \quad \forall b_\beta \in A_\beta.$$

We call $A_{\bar{0}}$ resp. $A_{\bar{1}}$ the "even" respectively the "odd" part of A.

We call A "super–commutative" or "$\mathbb{Z}_2$–graded–commutative" if

$$a_\alpha \cdot b_\beta = (-1)^{\alpha \cdot \beta} b_\beta \cdot \alpha_\alpha \quad \forall\ a_\alpha \in A_\alpha,\ \forall\ b_\beta \in A_\beta,\ \forall\ \alpha, \beta \in \{\bar{0}, \bar{1}\} = \mathbb{Z}_2.$$

Lemma. *Let* $A := \Lambda(V^*)$, $A_{\bar{0}} := \bigoplus_{k \geq 0} \Lambda^{2k} V^*$ and $A_{\bar{1}} := \oplus_{l \geq 0} \Lambda^{2l+1} V^*$. *Then* $A = A_{\bar{0}} \oplus A_{\bar{1}}$, *and* $(A, +, \wedge)$ *is a super–commutative, associative, unital super algebra over* $\mathbb{K}$.

Proof. Exercise. $\square$

Definition. Let U and V be $\mathbb{K}$–vector spaces and let $T : U \to V$ be a $\mathbb{K}$–linear map. For all $k \in \mathbb{N}$ we define the "pullback operator" $T^* : \bigotimes^k V^* \to \bigotimes^k U^*$ by

$$(T^* m)(u_1, \ldots, u_k) := m(T(u_1), \ldots, T(u_k)) \quad \forall m \in \otimes^k V^*,\ \forall u_1, \ldots, u_k \in U.$$

Lemma. *Let* U *and* V *be* $\mathbb{K}$*–vector spaces,* $T : U \to V$ $\mathbb{K}$*–linear and* k *in* $\mathbb{N}$. *Then*

(i) $T^* : \bigotimes^k V^* \to \bigotimes^k U^*$ *is* $\mathbb{K}$*–linear.*

(ii) $T^* : \mathcal{T}(V^*) = \bigoplus_{k \geq 0} \bigotimes^k V^* \to \mathcal{T}(U^*) = \bigoplus_{k \geq 0} \bigotimes^k U^*$ *is a* $\mathbb{K}$*–algebra homomorphism.*

(iii) $T^*(\Lambda^k V^*) \subset \Lambda^k U^*$

(iv) $T^*(\alpha \wedge \beta) = (T^*\alpha) \wedge (T^*\beta) \quad \forall\ \alpha, \beta \in \Lambda(V^*)$.

(v) Let W *be a further* $\mathbb{K}$*–vector space and* $S : V \to W$ *a* $\mathbb{K}$*–linear map. Then we have*

$$(S \circ T)^* = T^* \circ S^*.$$

Proof. Exercise. $\square$

Bibliographical remarks. *A standard reference for multilinear algebra is of course [Gre]. Good books on manifolds as [AMR] also give a thorough account of the material needed here.*

1.2 Volume and orientation

Definition. Let V be a m–dimensional $\mathbb{K}$–vector space and $T : V \to V$ a $\mathbb{K}$–linear map. The "determinant of the map T" is the number $\det(T)$ in $\mathbb{K}$ defined by

$$T^*\Omega = \det(T) \cdot \Omega \quad \forall\, \Omega \in \Lambda^m V^*.$$

Lemma. *Let V be a m–dimensional $\mathbb{K}$–vector space and T,S $\mathbb{K}$–linear endomorphisms of V. Then*

(i) $\det(T \circ S) = \det(T) \cdot \det(S)$,

(ii) (*T is invertible if and only if* $\det(T) \neq 0$),

(iii) *if $V = \mathbb{K}^m$ and $T = T_A$ for A in* $\mathrm{Mat}(m \times m, \mathbb{K})$ *then*

$$\det(T_A) = \det(A).$$

Remarks. We recall that $T_A(x) = A{\cdot}x$ and that $\det(A) = \sum\limits_{\sigma \in S_m} \mathrm{sign}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{m\sigma(m)}$ for A in $\mathrm{Mat}(m \times m, \mathbb{K})$.

Proof of the lemma. Exercise. $\square$

Definition. Let V be a m–dimensional $\mathbb{K}$-vector space. A non–zero element Ω in $\Lambda^m V^*$ is called a "volume form" or a "volume element (on V)"

If $\mathbb{K} = \mathbb{R}$ then such a Ω is also called an "orientation form".

Remark. If $\dim_{\mathbb{K}} V = m$, then $\dim_{\mathbb{K}} \Lambda^m V^* = 1$. Thus two volume forms are proportional.

Definition. Let V be a m–dimensional $\mathbb{R}$–vector space.

(1) Two orientation forms Ω' and Ω'' on V are called "equivalent" if $\Omega'' = \lambda \cdot \Omega'$ for a λ in $\mathbb{R}^{>0}$. We write $\Omega' \sim \Omega''$ or $[\Omega'] = [\Omega'']$.

(2) An equivalence class $[\Omega]$ of orientation forms on V is called an "orientation (on V)".

(3) If $[\Omega]$ is an orientation on V, then the pair $(V, [\Omega])$ is called an "oriented ($\mathbb{R}$–)vector space".

(4) An ordered basis $\{v_1, \ldots, v_m\}$ of an oriented vector space $(V, [\Omega])$ is called "positively (respectively negatively) oriented" if

$$\Omega(v_1, \ldots, v_m) > 0 \quad (\text{respectively} < 0).$$

Remark. The notions in (4) of the preceding definition are well–defined.

Lemma. *Let V be a finite dimensional $\mathbb{R}$–vector space. Then*

(i) *V has exactly two orientations and*

(ii) *the choice of an ordered basis of V uniquely determines an orientation on V.*

Proof. Exercise. $\square$

Definition. Let V and W be finite dimensional $\mathbb{K}$–vector spaces with volume forms Ω_V and Ω_W, and let $T : V \to W$ be a $\mathbb{K}$–linear map.

(1) T is called "volume preserving" if $T^*(\Omega_W) = \Omega_V$.

(2) If $\mathbb{K} = \mathbb{R}$ and T is a vector space isomorphism we define the "pullback orientation" on V by

$$T^*[\Omega_W] = [T^*\Omega_W].$$

In this situation, the map T is called "orientation preserving" respectively "orientation reversing" if

$$T^*[\Omega_W] = [\Omega_V] \quad \text{respectively} \quad T^*[\Omega_W] = -[\Omega_V].$$

Remark. The pullback orientation in (2) is well–defined.

Lemma. *Let V and W be finite dimensional $\mathbb{K}$–vector spaces with volumes forms Ω_V and Ω_W, and let $T : V \to W$ be $\mathbb{K}$–linear. Then*

(i) *T is volume preserving implies that T is a $\mathbb{K}$–vector space isomorphism,*

(ii) *if $\mathbb{K} = \mathbb{R}$, then a volume preserving map also preserves the induced orientations, i.e., $T^*[\Omega_W] = [\Omega_V]$,*

(iii) *if $\mathbb{K} = \mathbb{R}$ and $T^*\Omega_W = \lambda \cdot \Omega_V$ with $\lambda \in \mathbb{R}\backslash\{0\}$ then T is a $\mathbb{R}$–vector space isomorphism.*

Proof. *Ad(i).* The equality $T^*(\Omega_W) = \Omega_V$ implies immediately that $\dim_{\mathbb{K}} W = \dim_{\mathbb{K}} V =: m$. Assuming that T is not an isomorphism we find a non–zero vector v_1 in the kernel of T. We thus find – by the theorem on the completion to a basis – vectors $v_2, \ldots, v_m$ in V such that $\{v_1, v_2, \ldots, v_m\}$ is an ordered basis of V. Thus we have

$$0 \neq \Omega_V(v_1, \ldots, v_m) = (T^*\Omega_W)(v_1, \ldots, v_m) = \Omega_W(T(v_1), \ldots, T(v_m)) = 0.$$

This contradicts our assumption and hence T is a $\mathbb{K}$–vector space isomorphism.

Ad(ii). We have $T^*[\Omega_W] = [T^*\Omega_W]$ by definition and thus $T^*[\Omega_W] = [\Omega_V]$ for a volume preserving map T.

Ad(iii). This assertion follows from the same arguments as the first assertion. $\square$

Exercise. Let $V = \mathbb{R}^2$, with volume form $\Omega = e_1^* \wedge e_2^*$ and associated orientation $[\Omega]$. Find (2×2)–matrices A, B such that

(a) the map $T_A : V \to V, x \mapsto A{\cdot}x$ is orientation preserving, but not volume preserving, and

(b) the map T_B is a vector space isomorphism, but not orientation preserving.

Digression on the elementary geometric volume

Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ be two linearly independent vectors in $\mathbb{R}^2$ and let

$$P(x,y) = \{\lambda x + \mu y \in \mathbb{R}^2 \mid 0 \le \lambda, \mu \le 1\}$$

be the oriented parallelogram generated by x and y.

The volume $\mathrm{Vol}(P(x,y))$ of $P(x,y)$ is given by $\pm(h \cdot \|x\|)$, where h is the length of the projection of y orthogonal to x and the sign is determined by the elementary geometric definition of orientation (see below).

Let J be the anti-clockwise rotation by $\pi/2$, then it follows that $J\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$ and

$$h = \|\mathrm{pr}_{Jx} y\| = \frac{< Jx, y >}{\|Jx\|^2} Jx = \frac{|x_1y_2 - x_2y_1|}{\sqrt{x_1^2 + x_2^2}},$$

and thus

$$|\mathrm{Vol}(P(x,y))| = h \cdot \|x\| = |x_1y_2 - x_2y_1|.$$

An ordered pair (x,y) of two linearly independant vectors x, y in $\mathbb{R}^2$ is called "positively oriented in the elementary geometric sense" if and only if the anti-clockwise oriented angle from x to y is smaller than π. This is equivalent to saying that the (unoriented!) angle between Jx and y is at most $\pi/2$, i.e., $(\cos \sphericalangle(Jx,y)) = \frac{x_1y_2 - x_2y_1}{\|x\|\cdot\|y\|} > 0$.

Thus, the "oriented elementary geometric volume" $\mathrm{Vol}(P(x,y))$ of $P(x,y)$ is given by $x_1y_2 - x_2y_1$.

Exercise. Show that

$$\mathrm{Vol}(P(x,y)) = e_1^* \wedge e_2^*(x,y).$$

Let now $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$ be three linearly independent vectors in $\mathbb{R}^3$ and $P(x,y,z)$ the parallelotope generated by them. The unoriented volume of $P(x,y,z)$ is given by the following formula:

$$|\mathrm{Vol}(P(x,y,z))| = h \cdot a,$$

where $a = |\mathrm{Vol}(x,y)|$ and $h = \|\mathrm{pr}_{x \times y}(z)\|$ is the height, i.e. the length of the projection of z perpendicular to the plane generated by x and y.

The signs of the "cross–product" $x \times y$ as well as of the oriented volume of $P(x,y,z)$ are fixed by the "the right hand rule" and we find

$$\mathrm{Vol}(P(x,y,z)) = z_1(x_2y_3 - x_3y_2) + z_2(x_3y_1 - x_1y_3) + z_3(x_1y_2 - x_2y_1).$$

Exercise. Show that

$$\mathrm{Vol}(P(x,y,z)) = e_1^* \wedge e_2^* \wedge e_3^*(x,y,z).$$

Exercise. Let A be in $\mathrm{Mat}(3 \times 3, \mathbb{R})$, $T = T_A$ the associated endomorphism of $\mathbb{R}^3$, and x, y, z three linearly independent vectors in $\mathbb{R}^3$. Show that

$$\det(A) = \frac{\mathrm{Vol}(T(P(x,y,z)))}{\mathrm{Vol}(P(x,y,z))}.$$

Bibliographical remarks. *See the remarks at the end of the preceding section.*

1.3 Symplectic vector spaces

Definition (1) A symmetric or anti–symmetric bilinear form B on a vector space V is called "degenerate" if there exists a v in $V \backslash \{0\}$ such that $B(v, v') = 0$ for all v' in V. Otherwise B is called "non–degenerate".

(2) A non–degenerate anti–symmetric bilinear form on a vector space V is called a "symplectic form (on V)".

(3) A pair (V, ω) consisting of a vector space V and a symplectic form ω on V is called a "symplectic vector space".

Examples.

(1) Let W be a finite dimensional vector space and $V = W^* \oplus W$. Then the two–form ω defined by

$$\omega(v,v') = \omega((\varphi, w), (\varphi', w')) = \varphi(w') - \varphi'(w) \quad \forall\ v = (\varphi, w), v' = (\varphi', w') \in V$$

is symplectic.

(2) Let $V = \mathbb{K}^{2n}$ with canonical basis $\{e_1, ..., e_{2n}\}$ and v, v' in V be described by $v = \sum\limits_{j=1}^{n}(x_je_j + y_je_{n+j})$, $v' = \sum\limits_{j=1}^{n}(x'_je_j + y'_je_{n+j})$. Then

$$\omega(v,v') = \sum_{j=1}^{n}(x_j \cdot y'_j - y_j \cdot x'_j)$$

defines a symplectic form on V.

Denoting the identity matrix in $\mathrm{Mat}(n \times n, \mathbb{K})$ by $\mathbb{E}_n$ and the matrix $\begin{pmatrix} O & \mathbb{E}_n \\ -\mathbb{E}_n & 0 \end{pmatrix}$ in $\mathrm{Mat}(2n \times 2n, \mathbb{K})$ by $\mathbb{J}_n$, we can rewrite the above form as follows:

$$\omega(v, v') = {}^t v \cdot \mathbb{J}_n \cdot v'.$$

Yet another description is given by

$$\omega = \sum_{j=1}^{n} e_j^* \wedge e_{n+j}^*.$$

(3) *An infinite dimensional example*

Let $V = \mathcal{S}(\mathbb{R})$ and

$$\omega(f, g) := \int_{\mathbb{R}} f(t) \frac{dg(t)}{dt} dt \quad \forall\, f, g \in V.$$

It is easy to check that ω is an anti–symmetric bilinear form on V that is non–degenerate in the above defined sense.

Lemma. *Let ω be an antisymmetric bilinear form on a finite dimensional vector space V. Then the following are equivalent*:

(i) *ω is symplectic, i.e. ω is non–degenerate, and*

(ii) *the map $\omega^\flat : V \to V^*$, $\omega^\flat(v)(v') := \omega(v, v')$ is a vector space isomorphism.*

Remark. If V is infinite dimensional, ω should be continuous in an appropriate sense. Furthermore the above lemma is wrong in infinite dimensions, and fulfilling the assertion (i) respectively (ii) is referred to as "weak" respectively "strong" non–degeneracy of a (continuous) anti–symmetric bilinear form.

Proof of the lemma. Let us first observe that $\omega^\flat$ is a linear map for any bilinear form ω.

If $\omega^\flat$ is an isomorphism then

$$\{0\} = \ker \omega^\flat := \{v \in V | \omega(v, v') = 0 \quad \forall v' \in V\}$$

and ω is non–degenerate. Thus (ii) implies (i).

On the other hand (i) implies that $\omega^\flat$ is injective and in finite dimensions this assures us that $\omega^\flat : V \to V^*$ is an isomorphism. $\square$

Definition. Let ω be an anti–symmetric bilinear form on a vector space V and $M \subset V$ a subset. Then the "ω–annihilator of M" is defined as follows

$$M^\angle = \{v \in V \mid \omega(u, v) = 0 \quad \forall\, u \in M\}.$$

Notably in the case that $M = U$ is a linear subspace of V, we call $U^\angle$ the "skew–complement of U(in V)" or the "ω–complement of U(in V)".

Lemma. *Let ω be an anti–symmetric bilinear form on a $\mathbb{K}$–vector space V, and let M and M' be subsets of V. Then*

(i) $M \subset M'$ implies $(M')^{\angle} \subset M^{\angle}$,

(ii) $M^{\angle}$ is a linear subspace of V,

(iii) $M^{\angle} = \{((M))_{\mathbb{K}}\}^{\angle}$,

(iv) $M^{\angle\angle} = ((M))_{\mathbb{K}}$ if ω is symplectic,

(v) $(M \cup M')^{\angle} = M^{\angle} \cap (M')^{\angle}$,

(vi) $(M \cap M')^{\angle} = M^{\angle} + (M')^{\angle}$.

(vii) Let now $M = U$ be a linear subspace of V, then $U \cap U^{\angle} = \{u \in U | \omega(u, u') = 0$ $\forall\ u' \in U\}$ and on $U/(U \cap U^{\angle})$ one has the following canonical symplectic form

$$\omega_{\text{red}}([u_1], [u_2]) = \omega(u_1, u_2) \quad \text{for all} \quad u_1, u_2 \in U.$$

(viii) *Let V be of finite dimension, ω be non–degenerate and U a linear subspace of V, then*

$$\dim_{\mathbb{K}} U + \dim_{\mathbb{K}} U^{\angle} = \dim_{\mathbb{K}} V$$

and $(U^{\angle})^{\angle} = U$.

Remarks. (1) For a subset M in a vector space V the expression $((M))_{\mathbb{K}}$ denotes the linear subspace generated by M.

(2) For two linear subspaces U_1 and U_2 of V the sum $U_1 + U_2$ denotes the linear subspace generated by $U_1 \cup U_2$.

(3) The symplectic vector space $(U/(U \cap U^{\angle}), \omega_{\text{red}})$ is called the "symplectic reduction of U". In the case that $U = V$ one has $V \cap V^{\angle} = V^{\angle}$ and this space is also called the "degeneration space of ω", $\ker \omega$.

Proof of the lemma.

Ad(i) – (vi). Exercise.

Ad(vii). Since $\omega|_{U\times U}$ vanishes if one of the arguments is in $U \cap U^{\angle}$ the function ω_{red} is well–defined on $U/_{U\cap U^{\angle}} \times U/_{U\cap U^{\angle}}$. A direct verification shows that ω_{red} inherits bilinearity and anti–symmetry from ω.

Let $[u]$ be in Ker ω_{red}, then $\omega(u, u') = 0$ for all u' in U, i.e. u is in $U \cap U^{\angle}$ and therefore $[u] = 0$. Thus ω_{red} is non–degenerate on $U/(U \cap U^{\angle})$.

Ad (vi). The equality $U^{\angle} = \bigcap_{u\in U} \omega^{\flat}(u)$ shows that the codimension of $U^{\angle}$ in V equals the dimension of the subspace $\omega^{\flat}(U)$ of V^*. Since $\omega^{\flat}$ is injective we find

$$\dim V - \dim U^{\angle} = \dim U.$$

Obviously we have $U \subset (U^{\angle})^{\angle}$ and since $\dim U = \dim(U^{\angle})^{\angle}$ we have equality. $\square$

Lemma. *Let ω be an anti–symmetric bilinear form on a finite dimensional vector space V and $\tilde{V}$ a vector space complement of $V^\angle = \ker\omega$ in V (i.e. $V^\angle \oplus \tilde{V} = V$). Then $\omega|_{\tilde{V}\times\tilde{V}}$ is non–degenerate.*

Proof. Let $\tilde{v}$ be in $\tilde{V}$ such that $\omega(\tilde{v},\tilde{w}) = 0$ for all $\tilde{w}$ in $\tilde{V}$. Since $V = V^\angle \oplus \tilde{V}$ it follows that $\tilde{v}$ is skew–orthogonal to V and thus $\tilde{v} \in V^\angle \cap \tilde{V} = \{0\}$, showing the assertion. □

Definition. Let ω be an anti–symmetric bilinear form on V and U a subspace of V. Then U is called a "symplectic subspace of V (or of (V,ω))" if and only if $\omega|_{U\times U}$ is non–degenerate.

Remark. The preceding lemma can now be restated as follows: any vector space complement of the degeneracy space $V^\angle = \ker\omega$ is a symplectic subspace of V.

Lemma. *Let (V,ω) be a finite dimensional symplectic vector space and $U \subset V$ a subspace. Then the following are equivalent:*

(i) U is a symplectic subspace of (V,ω)

(ii) $U \cap U^\angle = \{0\}$

(iii) $V = U \oplus U^\angle$

(iv) $U^\angle$ is a symplectic subspace of (V,w).

Proof. Exercise. □

Remark. Let us observe that if U is symplectic in (V,ω) then the direct sum decomposition $V = U \oplus U^\angle$ is a "symplectic direct sum", i.e. U and $U^\angle$ are symplectic subspaces and $\omega(u,v) = 0$ for u in U and v in $U^\angle$.

Theorem ("Normal form of linear symplectic forms").

Let (V,ω) be a symplectic $\mathbb{K}$–vector space of finite dimension. Then there exists an ordered basis $\{e_1,\ldots,e_{2n}\}$ of V such that

$$\omega = \sum_{j=1}^{n} e_j^* \wedge e_{n+j}^*.$$

Proof. The dimension m of V is bigger then 0 since a zero–dimensional space does not carry a non–null two form.

Let e_1 be in $V\backslash\{0\}$, then there is a vector f_1' such that $0 \neq \lambda_1 = \omega(e_1, f_1')$. Setting $f_1 = (1/\lambda_1)f_1'$ we have $\omega(e_1,f_1) = 1$.

We set $V_1 = (((e_1,f_1))_{\mathbb{K}})^\angle$ and observe that V_1 is a symplectic ω–orthogonal complement to $((e_1,f_1))_{\mathbb{K}}$ in V.

Iteration of the above process yields a linear independant family $\mathcal{B} = \{e_1, f_1 \ldots, e_n, f_n\}$ such that

$$\omega(e_j,f_k) = \delta_{j,k} \text{ and } \omega(e_j,e_k) = 0 = \omega(f_j,f_k) \quad \text{for all } j,k = 1,\ldots,n,$$

and $V_{n+1} = \mathcal{B}^\angle$ is zero or one–dimensional. Since V_{n+1} is the ω–complement of the symplectic subspace generated by $\mathcal{B}$ it is a symplectic subspace of V. Thus $V_{n+1} = \{0\}$ and

$\mathcal{B}$ is a basis of V. Obviously we have $\omega = \sum\limits_{j=1}^{n} e_j^* \wedge f_j^*$ and renaming f_j as e_{n+j} the assertion of the theorem follows. $\square$

Corollary 1. *A finite dimensional symplectic vector space is of even dimension.*

Proof. Obvious from the above theorem. $\square$

Definition. An ordered basis $\{e_1, \ldots, e_{2n}\}$ of a symplectic vector space (V, ω) such that $\omega = \sum\limits_{j=1}^{n} e_j^* \wedge e_{n+j}^*$ is called a "symplectic basis of (V, ω)".

Remark. The word "symplectic basis" is also used for a basis $\{e_1, \ldots, e_n, f_1, \ldots, f_n\}$ as in the proof of the last theorem.

Corollary 2. *Let (V, ω) and (V', ω') be two symplectic vector spaces having the same finite dimension. Then there exists a vector space isomorphism $T : V \to V'$ such that $T^*(\omega') = \omega$.*

Proof. Let $2n = \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} V'$ and $\{e_1, \ldots, e_{2n}\}$ respectively $\{e_1', \ldots, e_{2n}'\}$ be symplectic basis of V respectively of V'. Defining

$$T : V \to V', \; T\Big(\sum_{j=1}^{2n} x_j e_j\Big) = \sum_{j=1}^{2n} x_j e_j'$$

we find $T^*(\omega') = \omega$. $\square$

Remark. Corollary 2 shows that all finite dimensional symplectic vector spaces "look like" the Example (2) given at the beginning of the section.

Corallary 3. *Let V be a finite dimensional vector space and ω an anti–symmetric bilinear form on V. Then there exists a basis $\{e_1, \ldots, e_{2p}, e_{2p+1}, \ldots, e_m\}$ such that $\omega = \sum\limits_{j=1}^{p} e_j^* \wedge e_{p+j}^*$.*

Proof. Exercise. $\square$

Lemma. *Let V be a $\mathbb{K}$–vector space having finite dimension m and ω in $\Lambda^2 V^*$. Then ω is symplectic if an only if $\omega^{\left[\frac{m}{2}\right]}$ is a volume form on V.*

Proof. Assuming that ω is symplectic we know that $m = 2n$ and thus $\left[\frac{m}{2}\right] = n$. Using a symplectic basis we have $\omega = \sum\limits_{j=1}^{n} e_j^* \wedge e_{n+j}^*$ and thus

$$\omega^n = \underbrace{\omega \wedge \ldots \wedge \omega}_{n \text{ factors}} = C_n \cdot (e_1^* \wedge e_{n+1}^* \wedge \ldots \wedge e_n^* \wedge e_{2n}^*)$$

with $C_n = (n!)$. Thus $\omega^n \neq 0$.

Assuming now that ω is not symplectic there exists $v_1 \neq 0$ in $V^{\angle} = \ker \omega$.

By induction we find that $\omega^k(v_1, v_2', \ldots, v_{2k}') = 0$ for all $k \geq 1$ and for all $v_2', \ldots, v_{2k}'$ in V. Taking $v_2, \ldots, v_{2\left[\frac{m}{2}\right]}$ such that $\{v_1, \ldots, v_{2\left[\frac{m}{2}\right]}\}$ is a linearly independent family it follows that $\omega^{\left[\frac{m}{2}\right]}(v_1, \ldots, v_{2\left[\frac{m}{2}\right]}) = 0$ and thus $\omega^{\left[\frac{m}{2}\right]}$ can not be a volume form on V. $\square$

Corollary. *Let (V,ω) be a symplectic vector space of finite dimension $2n$ and let*

$$\Omega = \left(\frac{(-1)^{\frac{(n-1)n}{2}}}{n!}\right)\omega^n.$$

Then for all symplectic basis $\{e_1,\ldots,e_{2n}\}$ of V we have $\Omega = e_1^* \wedge e_2^* \wedge \ldots \wedge e_{2n-1}^* \wedge e_{2n}^*$.

Proof. Exercise. □

Remark. The above defined Ω is called the "canonical volume form on (V,ω)" and $[\Omega]$ the "canonical orientation on (V,ω)".

Bibliographical remarks. *Any of our preferred standard references on symplectic geometry as [Ar1], [Bry] and [GS] does some linear symplectic geometry.*

1.4 Linear symplectic geometry

Definition. Let (V,ω) be a symplectic vector space of finite dimension and W a linear subspace of V. We call W

(1) "symplectic" if $\omega|_W$ is symplectic.

(2) "isotropic" if $\omega|_W$ is zero.

(3) "coisotropic" if $W^\angle$ is isotropic.

(4) "Lagrangian" (or a "Lagrange subspace") if W is isotropic and coisotropic.

Lemma. *For a linear subspace W in a finite dimensional symplectic vector space one has:*

(i) W is symplectic if and only if $W^\angle \cap W = \{0\}$.

(ii) W is isotropic if and only if $W \subset W^\angle$.

(iii) W is coisotropic if and only if $W^\angle \subset W$.

(iv) W is Lagrangian if and only if $W^\angle = W$.

Proof. Exercise. □

Theorem ("Normal form of subspaces of symplectic vector spaces").

Let W be a linear subspace of a symplectic vector space (V,ω) of dimension $2n$. Let $d = d(W) = \dim_{\mathbb{K}} W$ and $\nu = \nu(W) = \dim_{\mathbb{K}} W \cap W^\angle$ and $2n' = d - \nu$. Then there exists a symplectic basis $\{e_1,\ldots,e_n,f_1,\ldots,f_n\}$ of (V,ω) such that $\{e_1,\ldots,e_{n'},e_{n'+1},\ldots,e_{n'+\nu},f_1,\ldots,f_{n'}\}$ is a basis of W.

Proof. *Step 1: isotropic case.*

We first show the assertion by induction over ν in this case:

if $\nu = 0$ the result is trivial and for $\nu = 1$ the result easily follows from the existence of a vector f such that $\omega^\flat(f)|_W \neq 0$ and the fact that thus $W \oplus ((f))_{\mathbb{K}}$ is a symplectic subspace of V.

Assuming now that the assertion is true for a $\nu \geq 1$ we consider an isotropic subspace W of dimension $\nu + 1$. Let e' be a non–zero element of W and f' a vector in V such that $\omega(e', f') = 1$. The space $W' = \ker(\omega^\flat(f')|_W) \subset W$ is an isotropic subspace of (V, ω) of dimension ν and is contained in the symplectic vector space $V' = (((e', f'))_{\mathbb{K}})^{\angle}$. Applying the induction hypothesis to W' in V' and setting $e_{\nu+1} = e'$ yields the assertion.

Step 2: general case.

Given any linear subspace W in (V, ω) there is a symplectic subspace $\tilde{W}$ of V such that $W = \tilde{W} \oplus (W \cap W^\angle)$. Since $\tilde{W}$ is symplectic, $V = \tilde{W} \oplus (\tilde{W})^\angle$ is a symplectic direct sum (i.e. $\tilde{W}$ and $(\tilde{W})^\angle$ are symplectic and they are ω–perpendicular) and by construction $W \cap W^\angle \subset (\tilde{W})^\angle$. Applying the result of the theorem on the normal form of a symplectic form to $\tilde{W}$ and the result of Step 1 to $W \cap W^\angle \subset (\tilde{W})^\angle$ and numbering the union of the two basis appropriately we get the result. □

Corollary. *Let (V, ω) be a finite dimensional symplectic vector space. Then*

(i) each isotropic subspace of V is contained in a Lagrangian subspace of V.

(ii) each coisotropic subspace of V contains a Lagrangian subspace.

(iii) for each Lagrangian subspace L there is vector space isomorphism $A_L : V \to L \oplus L^*$ such that $A_L|_L = Id_L$ and $(A_L)^*\omega_{L\oplus L^*} = \omega$,

where $\omega_{L\oplus L^*}$ is the canonical symplectic form defined in Example (1) in the beginning of Section 1.3.

Proof. Exercise. □

Definition. Let (V_1, ω_1) and (V_2, ω_2) be symplectic vector spaces. The "symplectic sum of (V_1, ω_1) and (V_2, ω_2)" is the vector space $V_1 \oplus V_2$ together with the following 2–tensor

$$(\omega_1 \oplus \omega_2)((v_1, v_2), (v_1', v_2')) := \omega_1(v_1, v_1') + \omega_2(v_2, v_2') \quad \forall\, v_1, v_1' \in V_1 \quad \forall v_2, v_2' \in V_2.$$

Remark. It is easily checked that $\omega_1 \oplus \omega_2$ is anti–symmetric and non–degenerate, i.e. symplectic.

Definition. Let (V_1, ω_1) and (V_2, ω_2) be symplectic $\mathbb{K}$–vector spaces and $T : V_1 \to V_2$ $\mathbb{K}$–linear. We call T "symplectic" if $T^*\omega_2 = \omega_1$.

Remarks. (1) A symplectic linear map is always injective.

(2) If $V_1 = V_2 = V$ and $\omega_1 = \omega_2 = \omega$ then the canonical volume Ω is preserved by all symplectic maps.

If furthermore $\mathbb{K} = \mathbb{R}$ then the induced orientation $[\Omega]$ is preserved as well.

Lemma. *An endomorphism T of a finite dimensional symplectic vector space (V, ω) is symplectic if and only if Γ_T, the graph of the map T, is Lagrangian in $(V \oplus V, \omega \oplus (-\omega))$.*

Proof. Exercise. □

Definitions. (1) Let (V,ω) be a finite dimensional symplectic $\mathbb{K}$–vector space. Then the "symplectic group of (V,ω)" is defined as follows:

$$Sp(V,\omega) = \{T \in GL(V) \mid T^*\omega = \omega\}.$$

(2) If $V = \mathbb{K}^{2n}$ and $\omega_0 := \sum_{j=1}^{n} e_j^* \wedge e_{n+j}^*$ then we set $Sp(n,\mathbb{K}) := Sp(\mathbb{K}^{2n},\omega_0)$.

Remarks. (1) Obviously we have $Sp(V,\omega) \subset SL(V) := \{T \in GL(V) | \det T = 1\}$ and thus $Sp(n,\mathbb{K}) \subset SL(2n,\mathbb{K}) := SL(\mathbb{K}^{2n})$.

(2) We have the identity $Sp(1,\mathbb{K}) = SL(2,\mathbb{K})$.

Proposition. *Let W and W' be linear subspaces of a finite dimensional symplectic vector space (V,ω). Then there exists an element T of $Sp(V,\omega)$ such that $T(W) = W'$ if and only if* ($\dim_{\mathbb{K}} W = \dim_{\mathbb{K}} W'$ *and* $\dim_{\mathbb{K}} W \cap W^{\angle} = \dim_{\mathbb{K}} W' \cap (W')^{\angle}$).

Proof. We leave it as an exercise that $T(W) = W'$ for a T in $Sp(V,\omega)$ implies that the "numerical invariants" d and ν (defined in the above theorem) are the same.

Let now W and W' be subspaces of (V,ω) and let $\{e_1,\ldots,f_n\}$ and $\{e_1',\ldots,f_n'\}$ be symplectic basis of (V,ω) as given by the theorem on the normal form of subspaces of a symplectic vector space.

If $d(W) = \dim_{\mathbb{K}} W = \dim_{\mathbb{K}} W' = d(W')$ and $\nu(W) = \dim_{\mathbb{K}} W \cap W^{\angle} = \dim_{\mathbb{K}} W' \cap (W')^{\angle} = \nu(W')$ then the symplectic map T defined by $T(e_j) = e_j', T(f_j) = f_j'$ (for $j = 1,\ldots,n$) fulfills $T(W) = W'$ (and is of course in $Sp(V,\omega)$). $\square$

Bibliographical remarks. *See the remarks at the end of Section 1.3.*

1.5 Complex structures on real symplectic vector spaces

Definition Let V be a real vector space and g a symmetric bilinear form on V.

(1) We call g "non–degenerate" or a "pseudo–Riemannian metric (on V)" if for all $v \neq 0$ in V there is a w in V such that $g(v,w) \neq 0$.

(2) We call g "positive definite" or a "Riemannian metric (on V)" if $g(v,v) > 0$ for all $v \neq 0$ in V.

(3) A pair (V,g) of a real vector space V and a pseudo–Riemannian respectively a Riemannian metric g on V is called a "pseudo–Riemannian" respectively "Riemannian vector space".

Examples. (1) Let $V = \mathbb{R}^m$ and $g(x,y) = {}^t x{\cdot}y$. This is called the "standard Riemannian metric on $\mathbb{R}^m$".

(2) Let $V = \mathbb{R}^d = \mathbb{R} \oplus \mathbb{R}^{d-1}$ and an element of V be described as follows: $x = x_0 e_0 + \sum_{j=1}^{d-1} x_j e_j$. Then the "standard Lorentz metric on $\mathbb{R}^d$" is the following pseudo-Riemannian metric: $g_L(x, y) = x_0 \cdot y_0 - {}^t(x') \cdot (y')$, where $x' = {}^t(x_1, \dots, x_{d-1}), y' = {}^t(y_1, \dots, y_{d-1})$.

(3) Let V be the space of real $(m \times n)$–matrices $(m, n \geq 1)$ and g be defined as follows:

$$g(A, B) = \text{ trace } ({}^t A \cdot B) \quad \forall\ A, B \in V = \text{ Mat } (m \times n, \mathbb{R}).$$

The bilinear form g is symmetric and positive definite.

Definitions. (1) Let (V, g) be a pseudo–Riemannian vector space and U a linear subspace. The "orthocomplement of U (in (V, g))" is the subspace

$$U^\perp = \{v \in V \mid g(u, v) = 0 \quad \forall\ u \in U\}.$$

(2) For a pseudo–Riemannian vector space (V, g) we set $O(V, g) = \{T \in GL(V) \mid T^* g = g\}$, the "orthogonal group of (V, g)" and $SO(V, g) = O(V, g) \cap SL(V)$, the "special orthogonal group of (V, g)".

Remark. If (V, g) is a Riemannian vector space and U a linear subspace of V, then $V = U \oplus U^\perp$. This does not necessarily hold true for pseudo–Riemannian vector spaces.

Definition. Let V be a real vector space and J a real endomorphism of V. Then J is called a "complex structure (on the real vector space V)" if $J^2 = -\mathrm{Id}_V$.

Remarks. (1) Let W be a complex vector space and $W_\mathbb{R}$ the "underlying real vector space" obtained by restriction to the real scalars $(\mathbb{R} \subset \mathbb{C})$. The multiplication by $i = \sqrt{-1}$ on the set $W_\mathbb{R}$ is a real linear endomorphism. Denoting it by J we obviously have $J^2 = -Id_{W_\mathbb{R}}$ so that J is a complex structure on $W_\mathbb{R}$.

(2) If J is a complex structure on a real vector space V then the following map defines a multiplication of elements of V with complex scalars:

$$\mathbb{C} \times V \to V,\ (a + ib, v) \mapsto av + bJv \quad \text{for} \quad a, b \in \mathbb{R} \quad \text{and} \quad v \in V.$$

(3) If a real finite dimensional vector space V carries a complex structure then its dimension is even.

(4) If J is a complex structure on V then $-J$ is a complex structure as well.

(5) Let $V = \mathbb{R}^{2n}$ and $\mathbb{J}_n = \begin{pmatrix} 0 & -\mathbb{E}_n \\ \mathbb{E}_n & 0 \end{pmatrix}$ (where $\mathbb{E}_n$ is the identity matrix of size $(n \times n)$). Then $\pm \mathbb{J}_n$ are complex structures on V.

(6) If J is a complex structure on V we set $GL(V, J) = \{T \in GL_\mathbb{R}(V) \mid T \circ J = J \circ T\}$.

Exercise. Let $W = \mathbb{C}^n$ and $\Psi : W \to \mathbb{R}^{2n}$ be defined by $\Psi({}^t(z_1, \dots, z_n)) = {}^t(Re(z_1), \dots, Re(z_n), Im(z_1), \dots, Im(z_n))$. Then Ψ is a linear isomorphism (over the reals!) between $W_\mathbb{R}$ and $\mathbb{R}^{2n}$. Furthermore the complex structure J on $W_\mathbb{R}$ induced by multiplication by i on W fulfills: $\Psi \circ J = \mathbb{J}_n \circ \Psi$ (compare Remark (5) above for the definition of $\mathbb{J}_n$).

Finally we have $\Psi \circ C = C_{\mathbb{R}} \circ \Psi$, where for $C = M + iB$ in Mat $(n \times n, \mathbb{C})$ and A, B in Mat $(n \times n, \mathbb{R})$ the matrix $C_{\mathbb{R}}$ is given by $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$ in Mat $(2n \times 2n, \mathbb{R})$. Thus $C \mapsto C_{\mathbb{R}}$ defines a group isomorphism $GL(n, \mathbb{C}) = GL_{\mathbb{C}}(W) \to GL(\mathbb{R}^{2n}, \mathbb{J}_n)$.

Definitions. Let J be a complex structure on a real vector space V and W a linear subspace of V. Then W is called

(1) "complex" if $J(W) = W$, and

(2) "totally real" if $J(W) \cap W = \{0\}$.

Definitions. (1) Let (V, J) be a real vector space with a complex structure J. A "(pseudo–)Hermitian metric (on (V, J))" is a (pseudo–)Riemannian metric g on V such that $g(J(v), J(w)) = g(v, w)$ for all v, w in V.

(2) A "(pseudo–)Hermitian vector space" is a triple (V, J, g) consisting of a real vector space V, a complex structure J on V and a (pseudo–)Hermitian metric g on (V, J).

(3) The "unitary group" of a pseudo–Hermitian vector space (V, J, g) is given by $U(V, J, g) = \{T \in GL(V, J) \mid T^*g = g\} = GL(V, J) \cap O(V, g)$.

Proposition. *Let (V, J, g) be a pseudo–Hermitian vector space and let*

$$\omega(v, w) := g(J(v), w) \quad \forall\ v, w \in V.$$

Then ω is a symplectic form on V fulfilling $\omega(J(v), J(w)) = \omega(v, w)$ for all v and w in V.

Proof. Since g is bilinear and J is linear ω is a bilinear form on V. Since g is pseudo–Hermitian ω is anti–symmetric and $\omega(J(v), J(w)) = \omega(v, w)\ \forall\ v, w \in V$.

Given $v \neq 0$ there is a u in V such that $g(v, u) \neq 0$ since g is nondegenerate. Setting $w = J(u)$ we have $\omega(v, w) \neq 0$ and thus ω is nondegenerate. □

Remarks. (1) The symplectic structure ω associated by the above proposition to a (pseudo–)Hermitian vector space is called the "(pseudo–)Kähler form on (V, J, g)".

(2) For $V = \mathbb{R}^{2n}$, $J = \mathbb{J}_n$ and g the standard Riemannian structure on V, we observe that g is Hermitian on (V, J) and that the associated Kähler form is the standard symplectic form on $\mathbb{R}^{2n}$.

Lemma. *Let (V, J, g) be a pseudo–Hermitian vector space with associated pseudo–Kähler form ω. Then*

(i) $U^{\angle} = J(U^{\perp})$ for all linear subspaces U of V.

If g is Riemannian, i.e. positive definit, we have furthermore

(ii) all complex subspaces of V are symplectic, and

(iii) all isotropic subspaces of V are totally real.

Proof. Exercise. □

Remark. The last two assertions of the preceding lemma do not hold true if g is only non–degenerate.

Reversing the order of the data we have

Definitions. Let (V, ω) be a symplectic real vector space.

(1) A complex structure J on V is called "compatible with ω" if

$$g_J(v, w) := \omega(v, J(w)) \quad \forall\ v, w \quad \text{in } V$$

defines a Hermitian metric on (V, J).

(2) The set of all complex structures on V that are compatible with ω will be denoted by $\mathcal{J}_\omega(V)$.

Remarks. If a complex structure J on a symplectic vector space (V, ω) preserves ω then g_J, defined as above in (1), is pseudo–Hermitian, but not necessarily positive definite.

In this situation the pseudo–Kähler form of (V, J, g_J) is of course the original ω.

Denoting the set of Riemannian metrics on a vector space V by $\mathcal{R}(V)$ we have

Theorem. *Let (V, ω) be a real finite dimensional symplectic vector space. Then there is a real–analytic map*

$$\mathcal{R}(V) \xrightarrow{\Psi} \mathcal{J}_\omega(V), \quad g \mapsto J(g)$$

such that for each J in $\mathcal{J}_\omega(V)$ we have $\Psi(g_J) = J$, where g_J is defined in (1) of the last definition.

Corollary 1. *The set $\mathcal{J}_\omega(V)$ is not empty for a finite dimensional real symplectic vector space (V, ω).*

Corollary 2. *The topological space $\mathcal{J}_\omega(V) \subset End_{\mathbb{R}}(V)$ is smoothly contractible to a point, i.e. given a J_1 in $\mathcal{J}_\omega(V)$ there is a smooth map $H : [0,1] \times \mathcal{J}_\omega(V) \to \mathcal{J}_\omega(V)$ such that $H_0 = Id$, $H_1(J) \equiv J_1$ for all J in $\mathcal{J}_\omega(V)$, and $H_t(J_1) \equiv J_1$ for all t in $[0,1]$ $(H_t(J) := H(t, J))$.*

Proof of the theorem. Given a non–degenerate symmetric bilinear form g on V one has an isomorphism $g^\flat : V \to V^*, g^\flat(v)(w) := g(v, w)$. Thus for each g in $\mathcal{R}$(V) we have an isomorphism $A = A(g) = (g^\flat)^{-1} \circ \omega^\flat$ such that

$$\omega(v, w) = g(A(v), w) \quad \forall\ v, w \in V.$$

Since $\mathcal{R}(V)$ is open in the finite dimensional vector space $\mathcal{S}^2(V^*)$ it follows by the von Neumann series that A depends real–analytically on g.

Furthermore, it is easy to see that $g(v, A(w)) = -g(A(v), w)$, i.e. A is anti–selfadjoint with respect to g $(A^{*g} = -A)$. Thus $B = B(g) = A^{*g} \circ A = -A^2$ is self–adjoint and positive with respect to g (i.e. $g(B(v), v) > 0$ for $v \neq 0$).

It follows that B has a "square root" $\sqrt{B}$: a selfadjoint and positive operator such that $(\sqrt{B}) \circ (\sqrt{B}) = B$. Since the square root is given by the series of the square root function on $\mathbb{R}^+$ applied to a positive endomorphism we conclude that $B \mapsto \sqrt{B}$ is a real–analytic map. We define $J = J(g) = \Psi(g)$ by $\frac{A}{\sqrt{-A^2}} = A \circ (\sqrt{-A^2})^{-1} = (\sqrt{-A^2})^{-1} \circ A$

(A commutes with $\sqrt{-A^2}$). It follows that $J^2 = -\mathrm{Id}_V$, $J \circ A = A \circ J, J^{*g} = -J$ and $\omega(J(v), J(w)) = \omega(v, w) \quad \forall\ v, w \in V$. A direct calculation then shows that the pseudo–Hermitian metric $g_J\ (v, w) = \omega(v, J(w))$ is positive definite, i.e. $J = \Psi(g)$ is in $\mathcal{J}_\omega(V)$. If now $g = g_J$ for a J in $\mathcal{J}_\omega(V)$ then $g_J(A(v), w) = \omega(v, w) = g_J(J(v), w)$ and thus $A = J$. It obviously follows that $\frac{A}{\sqrt{-A^2}} = J$ and thus $\Psi(g_J) = J$ for all J in $\mathcal{J}_\omega(V)$. $\square$

Proof of the corollaries. Since $\mathcal{R}(V)$ is open and non–empty in $\mathcal{S}^2(V^*)$, the first corollary is obviously true.

Let now g_1 be an arbitrary element of $\mathcal{R}(V)$ then the map

$$K : [0,1] \times \mathcal{R}(V) \to \mathcal{R}(V), K(t, g) = (1-t)g + tg_1$$

is smooth if we give $\mathcal{S}^2(V^*) \subset \otimes^2 V^*$ any vector space topology, for example by identifying $\mathcal{S}^2(V^*)$ with symmetric matrices of the appropriate size. Furthermore, with $K_t(g) := K(t, g)$ we have

$$K_0(g) = g \quad \text{and} \quad K_1(g) = g_1 \quad \forall\ g \in \mathcal{R}(V), \quad \text{and} \quad K_t(g_1) = g_1 \quad \forall\ t \in [0,1].$$

Thus K is a "smooth retraction from $\mathcal{R}(V)$ to g_1 in $\mathcal{R}(V)$".

Using the map $\Psi : \mathcal{R}(V) \to \mathcal{J}_\omega(V)$ fulfilling $\Psi(g_J) = J$ for all J in $\mathcal{J}_\omega(V)$ we construct a retraction of $\mathcal{J}_\omega(V)$ as follows: Let J_1 be in $\mathcal{J}_\omega(V)$ and $g_1 = g_{J_1}$ and K the above retraction from $\mathcal{R}(V)$ to g_1. We set $H(t, J) = \Psi(K(t, g_J))$, then $H : [0,1] \times \mathcal{J}_\omega(V) \to \mathcal{J}_\omega(V)$ is smooth and fulfills the assertion of Corollary 2. $\square$

Exercises. (1) Fill in the details concerning the real–analytic dependence of $J(g)$ in the variable g.

(2) Show that Ψ is not injective.

Bibliographical remarks. *Compare, e.g., [Bry], [McDS] and [Wei2] for complex structures on real symplectic vector spaces.*

2. Elementary differential topology

The notion of differentiable manifolds and bundles over them allows to give a rigorous and useful framework for e.g. the following questions in mathematical physics:

* what is a force field on a set given by constraints in a configuration space?

* what does the reduction of the number of degrees of freedom mean and how can one use it to solve Hamilton's equation?

* what is curved space–time and what is a tensor on it?

* what is a Lie group and how does the representation theory of Lie groups relate to the representation theory of Lie algebras?

Of course, we can not go into all these subjects, but we need the basics of the calculus on manifolds in order to define and study "symplectic manifolds".

2.1 Differentiable manifolds

Definition. Let $(M, \mathcal{T})$ be a topological space and m in $\mathbb{N}_0$. We call $(M, \mathcal{T})$ a "topological manifold of dimension m" if for each p in M there is an open set U containing p, an open set V in $\mathbb{R}^m$ and a homeomorphism $\varphi : U \to V$. An open covering $\mathcal{U} = \{U_\alpha | \alpha \in A\}$ of M (A an index set) together with homeomorphisms $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha) = V_\alpha$, with V_α open in $\mathbb{R}^m$ is called an "atlas of the topological manifold M" and is often denoted by $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) | \alpha \in A\}$. Each U_α is called a "chart domain" and $\varphi_\alpha : U_\alpha \to V_\alpha$ a "coordinate chart on M". Furthermore we call the m–tuple of functions ${}^t(x_1^\alpha, \ldots, x_m^\alpha)$ defined by $x_j^\alpha = pr_j \circ \varphi_\alpha : U_\alpha \to \mathbb{R}$ a "local coordinate system on M" (pr_j is the projection on the j–th coordinate from $\mathbb{R}^m$ to $\mathbb{R}$).

Remark. If $U_{\alpha\beta} = U_\alpha \cap U_\beta$ is not empty the map

$$\varphi_{\beta\alpha} := (\varphi_\beta \circ \varphi_\alpha^{-1})\Big|_{\varphi_\alpha(U_{\alpha\beta})} \ : \ \varphi_\alpha(U_{\alpha\beta}) \to \varphi_\beta(U_{\alpha\beta})$$

is a homeomorphism and is called a "change of coordinates (map)".

Examples. (1) Let U be an open, non–empty subset of $\mathbb{R}^m, U_0 = U, V_0 = U$ and $\varphi_0 = \mathrm{Id} : U_0 \to V_0$.

Then $\mathfrak{A} = \{(U_0, \varphi_0)\}$ is an atlas of the m–dimensional topological manifold U.

(2) Let $S^m = \{y \in \mathbb{R}^{m+1} \,|\, \|y\| = 1\} \subset \mathbb{R}^{m+1}$ ($\|y\|^2 = \sum\limits_{j=1}^{m+1} |y_j|^2$ here) with the subspace topology induced from the metric topology on the normed vector space $(\mathbb{R}^{m+1}, \| \ \|)$. We set $U_j^\pm = \{y \in S^m | \pm y_j > 0\}$ and

$$\varphi_j^\pm : U_j^\pm \to \mathbb{R}^m, \varphi_j^\pm(y) = {}^t(y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_{m+1}) \ .$$

Then $\varphi_j^\pm$ is a homeomorphism onto its image $V_j^\pm = \{x \in \mathbb{R}^m | \ \|x\| < 1\}$ and its inverse is given as follows

$$(\varphi_j^\pm)^{-1} : V_j^\pm \to U_j^\pm, x \mapsto {}^t(x_1, \ldots, x_{j-1}, \pm\sqrt{1 - \|x\|^2}, x_j, \ldots, x_m) \ .$$

The set $\{(U_j^\pm, \varphi_j^\pm) | j = 1, \ldots, m+1\}$ is a topological atlas and S^m is a topological manifold of dimension m.

Definition. A "topological manifold of finite dimension" is a disjoint union of topological manifolds M_j of dimension m_j such that the set of the m_j is bounded. A topological manifold of dimension m is also called a " pure–dimensional topological manifold".

Example. The set $\{x \in \mathbb{R}^2 | \, \|x\|^2 \in \mathbb{Z}\}$ is a topological manifold of finite dimension, that is not pure dimensional.

Lemma. *Let $(M, \mathcal{T})$ be a topological manifold of finite dimension and $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha)|\alpha \in A\}$ a topological atlas on M. Then a subset U of M is open if and only if $\varphi_\alpha(U \cap U_\alpha)$ is open in $V_\alpha = \varphi_\alpha(U_\alpha) \subset \mathbb{R}^{m(\alpha)}$ for all α in A. (Here $\varphi_\alpha : U_\alpha \to \mathbb{R}^{m(\alpha)}$ vor a $m(\alpha)$ in $\mathbb{N}_0$.)*

Proof. If U is open, so are $U \cap U_\alpha$ and $\varphi_\alpha(U \cap U_\alpha)$, since $\varphi_\alpha : U_\alpha \to V_\alpha$ is a homeomorphism.

If $\varphi_\alpha(U \cap U_\alpha)$ is open in V_α then $U \cap U_\alpha$ is open in U_α and therefore in M. It follows that $U = \bigcup_{\alpha \in A}(U \cap U_\alpha)$ is open in M. □

Corollary. *Let M be a set, $\{U_\alpha|\alpha \in A\}$ a covering of M, $\varphi_\alpha : U_\alpha \to V_\alpha$ bijective with V_α open in $\mathbb{R}^{m(\alpha)}$ for all α in A (with $m(\alpha)$ in $\mathbb{N}_0$ and $\{m(\alpha)|\alpha \in A\}$ bounded). Assume furthermore that $\varphi_\alpha(U_{\alpha\beta})$ is open in V_α and $\varphi_{\beta\alpha} = \varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_{\alpha\beta}) \to \varphi_\beta(U_{\alpha\beta})$ is a homeomorphism for all α and β in A. Then $\mathcal{T} = \{U \subset M | \varphi_\alpha(U \cap U_\alpha) \text{ is open } \forall\ \alpha \in A\}$ is a topology on M, $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha)|\alpha \in A\}$ a topological atlas on $(M, \mathcal{T})$ and M a finite dimensional topological manifold.*

Proof. A direct inspection shows that $\mathcal{T}$ is a topology. Since $\varphi_\alpha(U_{\alpha\beta})$ is open for all α and β, U_α is open in $(M, \mathcal{T})$ and it follows easily that $\mathfrak{A}$ is a topological atlas. Therefore $(M, \mathcal{T})$ is a topological manifold of finite dimension. □

Definition. A topological atlas $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha)|\alpha \in A\}$ on a finite dimensional topological manifold $(M, \mathcal{T})$ is called a "smooth atlas" or a "C^∞–atlas" if $\varphi_{\beta\alpha}$ is a smooth map whenever $U_{\alpha\beta}$ is not empty.

Remarks.

(1) Similarly one defines a "C^k–atlas" (for $k \in \mathbb{N}^{\geq 1}$) and a "real–analytic atlas" (or "C^ω–atlas"). We will not need these notions in this text.

(2) If the charts take value in $\mathbb{C}^{n(\alpha)} \cong \mathbb{R}^{2 \cdot n(\alpha)}$ and the coordinate changes $\varphi_{\beta\alpha}$ are holomorphic maps, the atlas is called "holomorphic" or "complex–analytic". Such an atlas always has an "underlying smooth atlas" defined by identifying $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ and considering the $\varphi_{\beta\alpha}$ as smooth(!) real mappings.

(3) Replacing the "local model" $\mathbb{R}^m$ by an infinite dimensional vector space with an appropriate topology and notion of differentiability we would arrive at the notion of an "infinite dimensional manifold".

(4) If we require the $\varphi_{\beta\alpha}$ to be smooth in the last corollary the conclusion is that the atlas $\mathfrak{A}$ is smooth.

Definitions. (1) Two smooth atlases on a topological space $(M, \mathcal{T})$ are called "equivalent" if their union is again a smooth atlas.

(2) A "differentiable or smooth structure" on a topological space $(M, \mathcal{T})$ is an equivalence class of smooth atlases.

(3) The union of all atlases in a differentiable structure is called the "maximal atlas of the differentiable structure".

(4) A "differentiable or smooth manifold (of finite dimension)" is a pair consisting of a topological space $(M,\mathcal{T})$ and a differentiable structure on it.

(5) A coordinate chart in the maximal atlas of the differentiable structure of a differentiable manifold is called an "admissible (coordinate) chart (on the differentiable manifold)".

Remark. Practically one usually constructs a smooth atlas with as few as possible charts and considers the differentiable structure and maximal atlas defined by it.

Examples.

(1) Open sets $U \subset \mathbb{R}^m$ with the atlas given in the preceding example (1) are smooth manifolds.

(2) The spheres S^m with the atlas given in the preceding example (2) are smooth manifolds.

(3) Let $(M,\mathcal{T})$ be a differentiable manifold with an atlas $\{(U_\alpha,\varphi_\alpha)|\alpha\in A\}$, and let Ω be an open subset of M. Then $\{(U_\alpha\cap\Omega,\varphi_\alpha|_{U_\alpha\cap\Omega})|\alpha\in A \text{ and } U_\alpha\cap\Omega\neq\emptyset\}$ is a smooth atlas on Ω with the subspace topology.

(4) Let M and N be smooth manifolds. Then the cartesian product $M\times N$ carries a natural topology, the "product topology", and a natural differentiable structure.

(5) Let M and N be smooth manifolds. Then the disjoint union $M\dot{\cup}N$ carries a natural topology and a natural differentiable structure.

(6) Let $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$ and $\mathbb{P}_m(\mathbb{K})=\mathbb{P}(\mathbb{K}^{m+1})$ the "projective m–space over $\mathbb{K}$", i.e. the set of 1–dimensional linear subspaces of $\mathbb{K}^{m+1}$. Then there is a natural projection

$$\pi:\mathbb{K}^{m+1}\backslash\{0\}\to\mathbb{P}_m(\mathbb{K}),\pi((z_0,z_1,\dots,z_m))=[z_0,z_1,\dots,z_m]=(((z_0,z_1,\dots,z_m)))_{\mathbb{K}}.$$

(We will write elements of $\mathbb{K}^{m+1}$ here in this example as line–vectors.) In words $\pi(z)$ is the line through z. We set $U_j=\{z_j\neq 0\}\subset\mathbb{P}_m(\mathbb{K})$ and

$$\varphi_j:U_j\to V_j=\mathbb{K}^m,\quad [z_0\dots,z_m]\mapsto\left(\frac{z_0}{z_j},\dots,\frac{z_{j-1}}{z_j},\frac{z_{j+1}}{z_j},\dots,\frac{z_m}{z_j}\right)$$

for $j=0,1,\dots,m$. (Observe that U_j and φ_j are well–defined, since $[z]=[z']$ if and only if $z'=\lambda{\cdot}z$ for a λ in $\mathbb{K}\backslash\{0\}$.) The inverse is given by $\varphi_j^{-1}(w_1,\dots,w_m)=[w_1,\dots,w_j,1,w_{j+1},\dots,w_m]$, the intersection $U_{jk}=\{z_j\neq 0 \text{ and } z_k\neq 0\}$ and the change of coordinates on $\varphi_j(U_{jk})=\{w\in\mathbb{K}^m\,|w_k\neq 0\}$ as follows (here for $j<k$):

$$\varphi_k\circ\varphi_j^{-1}:\varphi_j(U_{jk})\to\varphi_k(U_{jk})\subset\mathbb{K}^m,$$

$$(w_1,\dots,w_m)\mapsto\left(\frac{w_1}{w_k},\dots,\frac{w_j}{w_k},\frac{1}{w_k},\frac{w_{j+1}}{w_k},\dots,\frac{w_{k-1}}{w_k},\frac{w_{k+1}}{w_k},\dots,\frac{w_m}{w_k}\right).$$

It follows that the atlas $\{(U_j,\varphi_j)|j=0,1,\dots,m\}$ is smooth for $\mathbb{K}=\mathbb{R}$ and complex–analytic for $\mathbb{K}=\mathbb{C}$.

(7) Let H be a N–dimensional real vector space with a Riemannian metric or a N–dimensional complex vector space with a Hermitian metric (on the underlying real vector space $H_{\mathbb{R}}$ with its natural complex structure). We call the set $G_k(H)=\{W\subset H \mid W$ is

a $\mathbb{K}$–linear subspace of dimension $k\}$ for k in $\{1, \ldots, N-1\}$ the "Grassmann manifold of k–planes in H". Let us define $U(W) := \{\Gamma_T \subset H \mid T : W \to W^\perp$ is $\mathbb{K}$–linear$\}$ and observe that $\{U(W) \mid W \in G_k(H)\}$ covers $G_k(H)$. Furthermore we set

$$\varphi_W : U(W) \to V(W) = \operatorname{Hom}_{\mathbb{K}}(W, W^\perp), \varphi_W(\Gamma_T) = T\,.$$

Since $\operatorname{Hom}_{\mathbb{K}}(W, W^\perp)$ has the natural norm $\|T\|_\infty := \sup\{\|T(w)\| \mid w \in W \text{ and } \|w\| = 1\}$ (here $\|\ \|$ is the norm induced by the Riemannian respectively Hermitian structure on H) and since all norms are equivalent on a finite dimensional vector space, we can use any basis of W and $W^\perp$ to identify $\operatorname{Hom}_{\mathbb{K}}(W, W^\perp)$ linearly and homeomorphically with Mat $((N-k)\times k, \mathbb{K}) \cong \mathbb{K}^{(N-k)\cdot k}$. This way we can consider φ_W as a bijection taking values in $\mathbb{K}^{(N-k)\cdot h}$, i.e., a chart.

Let now $W = \Gamma_{T_1}$ in $U(W_1)$ and $W = \Gamma_{T_2}$ in $U(W_2)$ for $W_1 \neq W_2$. Writing the identity of H as a map $W_1 \oplus W_1^\perp \to W_2 \oplus W_2^\perp$ we get a matrix of linear operators $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $a : W_1 \to W_2$, $b : W_1^\perp \to W_2$ etc. Since $W = \operatorname{Im} \begin{pmatrix} \mathrm{Id}_{W_1} \\ T_1 \end{pmatrix} = \{w_1 + T_1(w_1) |\ w_1 \in W\} = \operatorname{Im} \begin{pmatrix} \mathrm{Id}_{W_2} \\ T_2 \end{pmatrix}$ there is a $\mathbb{K}$–linear isomorphism $q : W_1 \to W_2$ such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \circ \begin{pmatrix} \mathrm{Id}_{W_1} \\ T_1 \end{pmatrix} = \begin{pmatrix} \mathrm{Id}_{W_2} \\ T_2 \end{pmatrix} \circ q : W_1 \to H = W_2 \oplus W_2^\perp.$$

Thus $T_2 = (c + d \circ T_1) \circ (a + b \circ T_1)^{-1} = (\varphi_{W_2} \circ (\varphi_{W_1})^{-1})(T_1)$ and the coordinate changes $\varphi_{W_2} \circ (\varphi_{W_1})^{-1}$ are "algebraic" maps defined on $\varphi_{W_1}(U(W_1) \cap U(W_2)) = \{T_1 \in \operatorname{Hom}_{\mathbb{K}}(W_1, W_1^\perp) | a + b \circ T_1$ is invertible $\}$. Hence $G_k(H)$ is a smooth manifold for $\mathbb{K} = \mathbb{R}$ and a complex–analytic manifold for $\mathbb{K} = \mathbb{C}$.

Exercises. Fill in the details in proving that the preceding examples (1) – (7) yield smooth (respectively complex–analytic) manifolds.

Definition. Let M and N be differentiable manifolds and let $F : M \to N$ be a continuous map. Then F is called "smooth" if for all p in M and for all admissible charts $\varphi : U = U(p) \to V$ near p and $\psi : U' = U'(F(p)) \to V'$ near $F(p)$ the maps

$$\psi \circ F \circ \varphi^{-1} : \varphi(U \cap F^{-1}(U')) \to V'$$

are smooth.

Remarks.

(1) It is of course enough to check that $\psi_\beta \circ F \circ \varphi_\alpha^{-1}$ is smooth for an admissible atlas $\{(U_\alpha, \varphi_\alpha) | \alpha \in A\}$ of M and an admissible atlas $\{(U'_\beta, \psi_\beta) | \beta \in B\}$ of N. (Of course $\psi_\beta \circ F \circ \varphi_\alpha^{-1}$ must be restricted to the open set $\varphi_\alpha(U_\alpha \cap F^{-1}(U'_\beta))$ and there is nothing to check if this set is empty.)

(2) Let us observe that this definition does not make sense if the coordinate changes are not smooth maps: this is the very reason why we consider smooth atlases. On the other

hand with the above definition every chart $\varphi : U \to V$ and its inverse $\varphi^{-1} : V \to U$ are now smooth maps.

(3) Obviously the composition of two smooth maps is again smooth.

Examples.

(1) $F : \mathbb{R} \to S^1, F(t) = (\cos t, \sin t)$.

(2) $F : S^{2n+1} \to \mathbb{P}_n(\mathbb{C}), F((z_0, \ldots, z_n)) = \pi((z_0, \ldots, z_n)) = [z_0, \ldots, z_n]$.

(3) Let $\psi : S^2 \backslash \{^t(0,0,1)\} \to \mathbb{R}^2$ be given by the stereographic projection from the 2–sphere to the plane $\{x_3 = 0\} \subset \mathbb{R}^3$, and $\varphi_0 : \{[z_0, z_1] \in \mathbb{P}_1(\mathbb{C}) | z_0 \neq 0\} = U_0 \to \mathbb{C}$ the usual chart $[z_0, z_1] \mapsto \frac{z_1}{z_0}$. Then the map $F : \mathbb{P}_1(\mathbb{C}) \to S^2$ defined by $F([0,1]) = {}^t(0,0,1)$ and $F|_{U_0} = \psi^{-1} \circ R \circ \varphi_0$, where $R : \mathbb{C} \to \mathbb{R}^2, w \mapsto {}^t(Re(w), Im(w))$, is a smooth map from the smooth (real!) manifold underlying the complex manifold $\mathbb{P}_1(\mathbb{C})$ to the 2–sphere.

Definition. Let M and N be smooth manifolds and $F : M \to N$ a smooth map. Then F is called a "diffeomorphism" if F is bijective and F^{-1} is smooth.

Examples. (1) Let A be in $GL(m, \mathbb{R})$ and $F = T_A : \mathbb{R}^m \to \mathbb{R}^m$. Then F is a diffeomorphism.

(2) The map $F : \mathbb{P}_1(\mathbb{C}) \to S^2$ defined in the preceding Example (3) of smooth maps is a diffeomorphism.

Definition. Let M be a smooth manifold of finite dimension and N a subset of M. We call N a "closed submanifold of M" if N is closed in M and if for each p in N there is an admissible chart $U(p) \overset{\varphi}{\to} V \subset \mathbb{R}^m$ such that $\varphi(p) = 0$ and $\varphi(U(p) \cap N) = V \cap \{x_1 = \ldots = x_n = 0\}$ for a n in $\{0, \ldots, m\}$. (Note that m and n may depend on the point p in N.)

Examples.

(1) Let $f_1, \ldots, f_k$ be smooth functions on $\mathbb{R}^m (1 \leq k \leq m)$ and $c_1, \ldots, c_k$ real numbers such that the gradients $\nabla f_1(p), \ldots, \nabla f_k(p)$ are linearly independant for all p in the common level set $\{f_1 = c_1, \ldots, f_k = c_k\} \subset \mathbb{R}^m$. Then this level set is a closed $(m-k)$–dimensional submanifold of $\mathbb{R}^m$.

(2) Let $\mathcal{L} = \{W \in G_2(\mathbb{R}^4) | W$ is Lagrangian $\}$ (with respect to the standard symplectic form $\omega = \epsilon_1^* \wedge \epsilon_3^* + \epsilon_2^* \wedge \epsilon_4^*$, where $\{\epsilon_1, \ldots, \epsilon_4\}$ denotes the canonical basis of $\mathbb{R}^4$ here. Then $\mathcal{L}$ is a closed submanifold of $G_2(\mathbb{R}^4)$.

In closing this subsection we would like to slightly restrict our class of manifolds.

Let us first recall that a topological space $(M, \mathcal{T})$ is called a "Hausdorff space" or simply "Hausdorff" if for distinct points p and q in M there are open sets U and V in M such that $p \in U, q \in V$ and $U \cap V = \emptyset$.

Examples. (1) Let Ω be open in $\mathbb{R}^m$ with the subspace topology induced from the usual metric topology on $\mathbb{R}^m$ and let p_1 and p_2 be distinct points in Ω. Then there exists ϵ_1 and ϵ_2 in $\mathbb{R}^{>0}$ such that $B_{\epsilon_k}(p_k) = \{x \in \mathbb{R}^m | \quad \|x - p_k\| < \epsilon_k\}$ are contained in Ω and disjoint, i.e. Ω is Hausdorff.

(2) Let $M = (\mathbb{R}\backslash\{0\}) \dot{\cup} \{p_1, p_2\}$ and let $\mathcal{T}$ be defined as follows: a subset $U' \subset M$ contained in $\mathbb{R}\backslash\{0\}$ is open if and only if it is open as a subset of $\mathbb{R}\backslash\{0\} \subset \mathbb{R}$ with its usual subspace

topology, and a subset $U \subset M$ containing p_1 or p_2 (or both of them) is open if and only if $(U \cap (\mathbb{R}\backslash\{0\})$ is open in $\mathbb{R}\backslash\{0\}$ and there is a $\epsilon > 0$ such that $B_\epsilon(0)\backslash\{0\}$ is contained in $U \cap (\mathbb{R}\backslash\{0\}))$. This defines a non–Hausdorff topological space that allows for a smooth atlas.

Let us also recall that a subset $\mathcal{B} \subset \mathcal{T}$ of the topology of a topological space $(M, \mathcal{T})$ is called a "basis of the topology" if for each U in $\mathcal{T}$ there is a family $\{U_\lambda | \lambda \in \Lambda\} \subset \mathcal{B}$ such that $\cup_{\lambda\in\Lambda} U_\lambda = U$. Furthermore we call a topological space "second countable" if its topology has a countable basis.

Examples.

(1) The set $\mathcal{B} = \{B_{1/n}(q) \mid n \in \mathbb{N}^{\geq 1}, q \in \mathbb{Q}^m\}$ is a countable basis of the usual topology of $\mathbb{R}^m$.

(2) A non–countable set M equipped with the discrete topology (i.e. $\mathcal{T}$ is the power set $\mathcal{P}(M)$) is not second–countable.

(3) There are non–second countable Hausdorff topological spaces that have a smooth atlas (see [Sp]).

In order to avoid lengthy statements in the sequel of the text – unless it is explicitely stated otherwise – we will assume that a "manifold" is always a finite dimensional, smooth, Hausdorff and second countable manifold.

On the other hand we will neither assume that a manifold is connected nor that it is pure–dimensional.

Bibliographical remarks. *Beside [Sp] quoted above we recommend [AMR],[Bo] and [Lan] for further reading on the foundations of differentiable manifolds.*

2.2 Lie groups and smooth actions

Definition. A "Lie group" is a manifold that carries a group structure such that the map

$$G \times G \to G\,,\ (g,h) \mapsto g \cdot h^{-1}$$

is smooth.

Remark. Equivalently one can ask for the smoothness of the following two maps:

$$I : G \to G, g \mapsto g^{-1} \text{ and } M : G \times G \to G, (g,h) \mapsto g \cdot h.$$

Examples.

(1) Let V be a real respectively complex vector space of finite dimension. Then the addition "+" of vectors makes $(V, +)$ into an abelian smooth respectively complex–analytic Lie group.

(2) Let Γ be a countable group. Then the discrete topology makes Γ a zero–dimensional manifold and it follows that Γ is a zero–dimensional Lie group.

(3) Considering S^1 as a subgroup of $\mathbb{C}\backslash\{0\}$ and S^3 as a subgroup of $\mathbb{H}\backslash\{0\}$ ($\mathbb{H}$ denotes the skew–field of quaternions) give rise to Lie group structures on these spheres.

(4) For $\mathbb{K} = \mathbb{R}$ and $\mathbb{C}$, the "general linear group" $GL(n, \mathbb{K})$ is open in Mat $(n \times n, \mathbb{K})$ and matrix multiplication induces a group structure on this set. The usual formulas for multiplication and the invese of a matrix show that $GL(n, \mathbb{R})$ is a smooth real Lie group and $GL(n, \mathbb{C})$ a "complex–analytic Lie group". Analogously $GL_{\mathbb{K}}(V)$ is a smooth respectively complex–analytic Lie group if V is a finite dimensional vector space over $\mathbb{K} = \mathbb{R}$ respectively over $\mathbb{K} = \mathbb{C}$.

(5) If V is a finite dimensional $\mathbb{K}$–vector space ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) then the "special linear group" $SL(V) = SL_{\mathbb{K}}(V) = \{T \in GL_{\mathbb{K}}(V) \mid \det(T) = 1\}$ is a closed submanifold of $GL_{\mathbb{K}}(V)$ and a Lie group. If V is furthermore equipped with a symplectic form ω then the some assertions hold for $Sp(V, \omega)$.

(6) The sets $O(n, \mathbb{K}) = \{A \in \text{ Mat}(n \times n, \mathbb{K}) \,|\, {}^tO \cdot O = \mathbb{E}_n\}$ are closed submanifolds of $\text{Mat}(n \times n, \mathbb{K})$ and Lie groups, the so–called "real (respectively complex) orthogonal groups".

(7) The sets $U(n) = \{A \in \text{ Mat } (n \times n, \mathbb{C}) | {}^t\bar{A} \cdot A = \mathbb{E}_n\}$ and $SU(n) = \{A \in U(n) | \det(A) = 1\}$ are closed submanifolds of $GL(n, \mathbb{C})$ (looked upon as a real manifold) and Lie groups.

Definition. Let G be a Lie group and H a subset of G. We call H a "closed Lie subgroup" if H is a closed submanifold of G and a subgroup of G.

Exercices. (1) Show that a closed Lie subgroup is a Lie group.

(2) Show that $SL(n, \mathbb{R})$ and $O(n, \mathbb{R})$ are closed Lie subgroups of $GL(n, \mathbb{R})$.

Definition. Let G_1 and G_2 be two Lie groups. A map $F : G_1 \to G_2$ is called a "Lie group homomorphism" if F is smooth and a group homomorphism. A Lie group homomorphism F is called a "Lie group isomorphism" if F is bijective and F^{-1} is a Lie group homomorphism. We call two Lie groups G_1 and G_2 "isomorphic (as Lie groups)" if there exists a Lie group isomorphism $F : G_1 \to G_2$.

Example. The following Lie groups are isomorphic: $S^1, U(1)$ and $SO(2, \mathbb{R}) = O(2, \mathbb{R}) \cap SL(2, \mathbb{R})$.

Definition. Let M be a manifold and G a Lie group. A smooth map $\vartheta : G \times M \to M$ is called a "smooth (left)–action (of G on M)" if the following two conditions hold

(i) $\vartheta(e, p) = p$ for all p in M (e denotes the neutral element of the group G),

(ii) $\vartheta(g, \vartheta(h, p)) = \vartheta(g{\cdot}h, p)$ for all g, h in G and all p in M.

Remarks. (1) The map $M \to M, p \mapsto \vartheta(g, p)$ for fixed g is often denoted by ϑ_g. Since $\vartheta_g \circ \vartheta_{g^{-1}} = \text{Id}_M$ it follows that ϑ_g is a diffeomorphism for each g in G. Another useful notation is $g{\cdot}p$ for $\vartheta_g(p)$.

(2) A right-action is defined similarly replacing condition (ii) by

(ii') $\vartheta(g, \vartheta(h, p)) = \vartheta(h{\cdot}g, p)$ for all g, h in G and all p in M.

Examples.

(1) $GL(n,\mathbb{R})\times\mathbb{R}^n\to\mathbb{R}^n, \vartheta(A,x)=T_A(x)=A\cdot x$.

(2) $GL(n,\mathbb{R})\times G_k(\mathbb{R}^n)\to G_k(\mathbb{R}^n), \vartheta(A,W)=T_A(W)$.

(3) $Sp(n,\mathbb{R})\times\mathcal{L}\to\mathcal{L}$, where $\mathcal{L}=\{W\subset G_n(\mathbb{R}^{2n})|W$ is Lagrangian$\}$ and ϑ as in (2).

(4) Let G be a Lie group and $M=G$, then $\vartheta^L : G\times M\to M, \vartheta^L(g)(h)=g\cdot h$ and $\vartheta^R : G\times M\to M, \vartheta^R(g)(h)=h{\cdot}g^{-1}$ are smooth transitive actions. (5) $\vartheta:\mathbb{Z}_2\times S^m\to S^m, \vartheta(\bar{0},x)=x, \vartheta(\bar{1},x)=-x$.

(6) $\vartheta:\mathbb{Z}^n\times\mathbb{R}^n\to\mathbb{R}^n, \vartheta(a,x)=x+a$.

(7) $\vartheta : GL(n+1,\mathbb{K})\times\mathbb{P}_n(\mathbb{K})\to\mathbb{P}_n(\mathbb{K}), \vartheta(A,[z])=[A\cdot z]$ for $\mathbb{K}=\mathbb{R}$ and $\mathbb{C}$. (The line generated by z is as usual denoted by $[z]$.)

Definition. Let $\vartheta : G\times M\to M$ be a smooth action.

(1) The action is called "transitive" if for each pair p,q in M there is at least one g in G such that $g\cdot p=q$.

(2) The "fixed point set of an element g in G" is the set $\{p\in M|g\cdot p=p\}$ and the "fixed point set of G" is the set $M^G=\bigcap\limits_{g\in G}\{p\in M|g\cdot p=p\}=\{p\in M|g\cdot p=p\ \forall g\in G\}$.

(3) The "G–orbit through p" is the set $\{g\cdot p|g\in G\}\subset M$ (for p a point in M).

(4) The "isotropy group (or stabilizer) of p (under the G–action)" is $G_p=\{g\in G|g{\cdot}p=p\}$.

(5) The "orbit space" is the space of equivalence classes $M/_\sim=\{[p]|p\in M\}$, where $p\sim q$ if and only if there exists a g in G such that $g\cdot p=q$. The map $M\to M/_\sim, p\mapsto[p]$ is called the "canonical projection" and is often denoted by π.

Exercice. Go through the examples of actions (1) – (7) above and determine the fixed point sets of the group, the G–orbits, the isotropy groups and (set–theoretically) the orbit spaces of these actions.

Definition. Let Γ be a discrete Lie group, M a manifold and $\vartheta:\Gamma\times M\to M$ a smooth action. We call the action "free" if the fixed point set is empty for all γ in $\Gamma\backslash\{e\}$. We call the action "properly discontinuous" if for all K compact in M the set $\{\gamma\in\Gamma|\gamma(K)\cap K\neq\emptyset\}$ is finite.

Proposition. *Let $\vartheta:\Gamma\times M\to M$ be a properly discontinuous and free action. Then the orbit space $M/_\sim$ carries a unique differentiable structure such that the projection $\pi : M\to M/_\sim$ is a local diffeomorphism.*

Remark. A smooth map $F : M\to N$ is a "local diffeomorphism" if for each p in M there is an open set U containing p such that $F|_U : U\to F(U)$ is a diffeomorphism.

Proof of the proposition.

Let $\mathcal{T}$ be the quotient topology on $M/_\sim$, i.e. U is in $\mathcal{T}$ if and only if $\pi^{-1}(U)$ is open in M. Since M is a manifold and Γ is countable it follows easily that $\mathcal{T}$ has a countable basis.

Let now x be in M and $W\xrightarrow{\varphi}V$ be a coordinate chart such that $\mathbb{B}_2(\varphi(x))\subset V$. Let $W_{1/n}=\varphi^{-1}(\mathbb{B}_{1/n}(\varphi(x))$ and $K_n=\overline{W_{1/n}}$ for $n\geq 1$. (The notation $\mathbb{B}_\epsilon(p)$ denotes the ϵ–ball around a point p in $\mathbb{R}^m$.) Then the K_n are compact and contained in W. Consider now a $\gamma \neq e$ such that $\gamma(K_1) \cap K_1 \neq \emptyset$. Since K_1 is compact and γ has no fixed–point it follows easily that there is a $N(\gamma) \geq 1$ such that $\gamma(K_n) \cap K_n = \emptyset$ for $n \geq N(\gamma)$. Since the set $\{\gamma \in \Gamma | \gamma(K_1) \cap K_1 \neq \emptyset\}$ is finite there is a $N \geq 1$ such that $K_N \cap \gamma(K_N) = \emptyset$ for all $\gamma \neq e$. Thus $\pi|_{W_{1/N}} : W_{1/N} \to \pi(W_{1/N})$ is a homeomorphism.

Starting with any atlas we can now construct an atlas $\{(W_x, \Psi_x) | x \in M\}$ such that $W_x \cap \gamma(W_x) = \emptyset$ for all $\gamma \neq e$ (and for all $x \in M$). Since $\pi(W_x) =: U_x$ is an open set in $M/_\sim$ containing $\pi(x)$ we have an open covering $\{U_x | x \in M\}$ for $M/_\sim$ and homeomorphisms

$$\varphi_x = \Psi_x \circ (\pi|_{W_x})^{-1} : U_x \to \Psi_x(W_x).$$

Let now $y \neq x$ such that $U_x \cap U_y$ contains a point $\pi(z)$ (with $z \in W_x$.) Then by construction of the W_p for p in M the connected component of $(\pi|_{W_x})^{-1}(U_x \cap U_y)$ containing z is in $W_x \cap \gamma(U_y)$ for a unique γ in Γ. It follows that the coordinate changes of the topological atlas $\{(U_x, \varphi_x) | x \in M\}$ of $M/_\sim$ are given by $\varphi_y \circ \varphi_x^{-1} = \Psi_y \circ \vartheta_{\gamma^{-1}} \circ \Psi_x^{-1}$ for appropriate γ in Γ, thus they are smooth maps.

Let now $\pi(x) \neq \pi(y)$ in $M/_\sim$, i.e. $\Gamma \cdot x \cap \Gamma \cdot y = \emptyset$ for x and y in M. Defining $K_n(x)$ and $K_n(y)$ as in the beginning of the proof and $K = K_{N_0}(x) \cup K_{N_0}(y)$ for N_0 sufficiently big such that $K_{N_0}(x) \cap K_{N_0}(y) = \emptyset$, we have only a finite set of $\gamma \neq e$ with $K \cap \gamma(K) \neq \emptyset$. Obviously this implies $K_{N_0}(x) \cap \gamma(K_{N_0}(y)) \neq \emptyset$ or $\gamma(K_{N_0}(x)) \cap K_{N_0}(y) \neq \emptyset$. Compactness of the $K_N(x)$ and $K_N(y)$ and the absence of fixed point for $\gamma \neq e$ imply then that for N sufficiently big $\left(\Gamma \cdot W_{1/N}(x)\right) \cap \left(\Gamma \cdot W_{1/N}(y)\right) = \emptyset$, i.e. $\pi(W_{1/N}(x))$ and $\pi(W_{1/N}(y))$ are disjoint open neighborhoods of $\pi(x)$ and $\pi(y)$. Thus the quotient topology on $M/\sim$ is Hausdorff. $\square$

Examples of properly discontinuous and free actions.

(1) $\vartheta : \mathbb{Z}_2 \times S^m \to S^m$ as in Example (5) of smooth actions. The orbit space with the induced differentiable structure is diffeomorphic to $\mathbb{P}_m(\mathbb{R})$

(2) Let Γ be the $\mathbb{Z}$–module in $\mathbb{R}^m$ generated by k $\mathbb{R}$–linearly vectors $v_1, \ldots, v_k$ in $\mathbb{R}^m$. Then the orbit space $\mathbb{R}^m/_\sim$ is diffeomorphic to $(S^1)^k \times \mathbb{R}^{m-k}$.

(3) A general recipe to produce free, properly discontinuous actions is the following:

let Γ be a subgroup of a Lie group G such that the subspace topology on Γ is the discrete topology. Then Γ is a discrete Lie group and acts by restricting to it the actions ϑ^L and ϑ^R on $M = G$. Let, for example, $\vartheta = \vartheta^L|_\Gamma : \Gamma \times M \to M$. Then for each γ in $\Gamma \backslash \{e\}$ the fixed point set $\{g \in M = G | \gamma \cdot g = g\}$ is empty.

Assuming now that the action ϑ is not properly discontinuous, then there is a compact set K in G and a sequence $\{\gamma_n | n \in \mathbb{N}^{\geq 1}\}$ in Γ such that $\gamma_n \neq \gamma_m$ for $n \neq m$ and $\gamma_n(K) \cap K \ni g_n = \gamma_n \cdot h_n$ with g_n and h_n in K for $n \geq 1$. Going to subsequences and relabelling there are g_0 and h_0 in K such that $g_n \to g_0$ and $h_n \to h_0$. Thus finally we have $\gamma_n \to g_0 \cdot h_0^{-1}$, a convergent sequence in Γ with $\gamma_n \neq \gamma_m$ for $n \neq m$. This contradicts the fact that Γ carries the discrete topology and thus our assumption was wrong and Γ acts properly discontinously by $\vartheta = \vartheta^L|_\Gamma$ on $M = G$.

The quotient $M/_\sim$ being the set $\{\Gamma \cdot g | g \in G\}$ of left restclasses, we denote it sometimes by Γ/G as in abstract group theory. The corresponding quotient for the action $\vartheta^R|_\Gamma$ is then of course denoted by G/Γ.

As an illustration we give some concrete examples.

(3.1) Let R be a unital ring and

$$G_R = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \middle| a, b, c \in R \right\}.$$

Then G_R is a group, and $G_\mathbb{R}$ respectively $G_\mathbb{C}$ is a real Lie group respectively a complex–analytic Lie group. Examples of discrete Lie subgroups are $G_\mathbb{Z} \subset G_\mathbb{R}$ and $G_{\mathbb{Z}[i]} \subset G_\mathbb{C}$, and $G_\mathbb{Z} \backslash G_\mathbb{R}$ respective $G_{\mathbb{Z}[i]} \backslash G_\mathbb{C}$ are compact smooth respective complex–analytic manifolds.

(3.2) Let R be as in (3.1) and

$$N_R = G_R \times R = \left\{ \left(\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, d \right) \middle| a, b, c, d \in R \right\},$$

the product of the group G_R, defined above in (3.1) with the abelian group $(R, +)$. Then $N_\mathbb{Z}$ is a discrete subgroup of $N_\mathbb{R}$, and $N_\mathbb{Z} \backslash N_\mathbb{R}$ and $N_\mathbb{R}/N_\mathbb{Z}$ are compact smooth real four–dimensional manifolds.

Bibliographical remarks. *A good reference for the elementary theory of quotients by discrete groups is [Bo]. A lot of material on differentiable actions of compact Lie groups can be found in the classic [Bre].*

2.3 Vector bundles

Definition. A "smooth (real respectively complex) vector bundle of rank r (over a manifold M)" is a manifold E together with a smooth projection $p : E \to M$ such that for each x in M the fiber $p^{-1}(x) =: E_x$ has the structure of a r–dimensional (real respectively complex) vector space and such that the following condition of "local triviality" is satisfied: for each x in M there is an open set U containing x and a diffeomorphism $\Psi_U : p^{-1}(U) \to U \times \mathbb{K}^r$ such that $pr_1 \circ \Psi_U = p$ and $\Psi_U | E_x : E_x \to \{x\} \times \mathbb{K}^r$ is a $\mathbb{K}$–vector space isomorphism for each $x \in U$. ($\mathbb{K}$ equals $\mathbb{R}$ or $\mathbb{C}$ here).

Remark. We will often speak of the "vector bundle $E \xrightarrow{p} M$" in order to have a short notation including the projection map p. Furthermore E will sometimes be called the "total space" and M the "base" of the vector bundle.

Examples. Let $\mathbb{K}$ be $\mathbb{R}$ or $\mathbb{C}$.

(1) The product manifold $M \times \mathbb{K}^r$ with $p = pr_1$ is a vector bundle, the so–called "trivial vector bundle of rank r over M".

(2) Let $\vartheta : \mathbb{Z} \times (\mathbb{R} \times \mathbb{R}) \to \mathbb{R} \times \mathbb{R}, \vartheta(n,(x,v)) = (x+n, (-1)^n \cdot v)$ and $E = (\mathbb{R} \times \mathbb{R})/_\sim \xrightarrow{p} \mathbb{R}/\mathbb{Z}, p([x,v]) = [x]$.

Then E is a vector bundle over $\mathbb{R}/\mathbb{Z} \cong S^1$.

(3) Let $H = \{([z],v) \in P_m(\mathbb{K}) \times \mathbb{K}^{m+1} | v \in ((z))_{\mathbb{K}}\}$. Then H is a closed submanifold of $\mathbb{P}_m(\mathbb{K}) \times \mathbb{K}^{m+1}$ and the fibers of the projection $p : H \to \mathbb{P}_m(\mathbb{K}), p([z],v) = [z]$ are the one–dimensional $\mathbb{K}$–vector spaces $p^{-1}([z]) = \{([z],v)|v \in ((z))_{\mathbb{K}}\}(\cong ((z))_{\mathbb{K}} \subset \mathbb{K}^{m+1})$.

"Local trivializations" are obtained as follows: let $U_j = \{[z] \in \mathbb{P}_m(\mathbb{K})|z_j \neq 0\}$ and $\chi_j : U_j \times \mathbb{K} \to p^{-1}(U_j)$ be defined by

$$\chi_j([z],\lambda) = \Big([z], \lambda \cdot (\frac{z_0}{z_j}, \ldots, \frac{z_{j-1}}{z_j}, 1, \frac{z_{j+1}}{z_j}, \ldots, \frac{z_m}{z_j})\Big).$$

It is easy to check that χ_j is a diffeomorphism, that is "fiber-wise linear" and thus $\Psi_j = \chi_j^{-1}$ is a local trivialization of $H \xrightarrow{p} \mathbb{P}_m(\mathbb{K})$ over U_j (for $j = 0, \ldots, m$).

Definitions. (1) Let $E \xrightarrow{p} M$ a vector bundle. A "smooth section of E" is a smooth map $s : M \to E$ such that $p \circ s = \mathrm{Id}_M$. The set of smooth sections of E form a vector space which is denoted by $\Gamma_{C^\infty}(M,E)$. The "zero–section" is defined by $s : M \to E, s(x) = 0_x$, where 0_x is the zero–element of the vector space $E_x = p^{-1}(x)$.

(2) Let $E_1 \xrightarrow{p_1} M_1$ and $E_2 \xrightarrow{p_2} M_2$ be vector bundles and let $f : M_1 \to M_2$ be a smooth map. A smooth map $F : E_1 \to E_2$ is called a "smooth vector bundle homomorphism (over f)" if $F((E_1)_x) \subset (E_2)_{f(x)}$ for all x in M_1 and $F|_{(E_1)_x} : (E_1)_x \to (E_2)_{f(x)}$ is a linear map for all x in M_1. (Sometimes we refer to these properties by saying that F is "fiber–preserving" and "fiber–wise linear".)

If f is a diffeomorphism and $F|_{(E_1)_x}$ is an isomorphism for all x in M the map F is called a "smooth vector bundle isomorphism (over f)".

(3) Two vector bundles over the same base manifold M are called "isomorphic" if there exists a smooth vector bundle isomorphism over $f = \mathrm{Id}_M$.

A vector bundle is called "trivializable" or shortly "trivial" if it is isomorphic to the trivial bundle of the same rank.

Remark. The vector bundle E of Example (2) is non-trivial.

Definition. Let $E \xrightarrow{p} M$ be a vector bundle of rank r and $\{U_\alpha|\alpha \in A\}$ an open covering of M such that $\Psi_\alpha : p^{-1}(U_\alpha) \to U_\alpha \times \mathbb{K}^r$ are local trivializations of E. Then the maps $g_{\alpha\beta} : U_{\alpha_\beta} = U_\alpha \cap U_\beta \to GL(\mathbb{K}^r)$ defined by

$$\Psi_\alpha \circ \Psi_\beta^{-1} : U_{\alpha\beta} \times \mathbb{K}^r \to U_{\alpha\beta} \times \mathbb{K}^r, \Psi_\alpha \circ \Psi_\beta^{-1}(x,v) = (x, g_{\alpha\beta}(x) \cdot v)$$

are called the "transition functions of E with respect to the local trivializations $\{U_\alpha|\alpha \in A\}$".

Remark. The family $\{g_{\alpha\beta}|\alpha,\beta \in A\}$ fulfills the following "cocycle identities":

(1) $g_{\alpha\alpha}(x) = 1$ for all x in U_α,

(2) $(g_{\alpha\beta}(x))^{-1} = g_{\beta\alpha}(x)$ for all x in $U_{\alpha\beta}$,

(3) $g_{\alpha\beta}(x) \cdot g_{\beta\gamma}(x) \cdot g_{\gamma\alpha}(x) = 1$ for all x in $U_{\alpha\beta\gamma} = U_\alpha \cap U_\beta \cap U_\gamma$.

Proposition. *Let M be a manifold. Then*

(i) If $\{U_\alpha | \alpha \in A\}$ is an open covering of M and $g_{\alpha\beta} : U_{\alpha\beta} \to GL(\mathbb{K}^r)$ is a family of maps fulfilling the cocycle identities given above then the set $\left(\dot{\bigcup}_{\alpha\in A}(U_\alpha \times \mathbb{K}^r)\right)\Big/\sim$, where (x, v) in $U_\beta \times \mathbb{K}^r$ and (y, w) in $U_\alpha \times \mathbb{K}^r$ are equivalent if and only if $x = y \in U_{\alpha\beta}$ and $w = g_{\alpha\beta}(x) \cdot v$, is a vector bundle of rank r over M.

(ii) If $E \xrightarrow{p} M$ is a vector bundle of rank r and if $\{g_{\alpha\beta} | \alpha, \beta \in A\}$ are its transition functions with respect to a trivializing open cover $\{U_\alpha | \alpha \in A\}$ then E is isomorphic as a vector bundle to $\left(\dot{\bigcup}_{\alpha\in A}(U_\alpha \times \mathbb{K}^r)\right)\Big/\sim$.

(iii) Given the situation of (ii) then a section s of E defines (and is defined by) a family of smooth maps $s_\alpha : U_\alpha \to \mathbb{K}^r \, (\alpha \in A)$ such that $s_\alpha(x) = g_{\alpha\beta}(x) \cdot s_\beta(x)$ for all x in $U_{\alpha\beta}$ and for all α, β in A.

Proof. Exercise. □

Remarks. (1) Using the transition functions we can apply pointwise the usual operations from multilinear algebra to construct new vector bundles from given ones. Let E and F be vectorbundles then $E \oplus F$ with fiber $(E \oplus F)_x = E_x \oplus F_x$ is called the "direct sum of the vector bundles E and F". Similarly one defines $E \otimes F$, $\mathrm{Hom}\,(E, F)$, and, if $F \subset E$ is a "subbundle", E/F. Furthermore we have the bundles E^* with fiber $(E^*)_x = (E_x)^*$ and $\otimes^k E^*, \mathcal{S}^k E^*$ and $\Lambda^k E^*$ for $k \in \mathbb{N}_0$.

(2) If $f : M \to N$ is a smooth map and $p : E \to N$ a vector bundle we have the "pullback of E under f" (or simply the "pullback bundle") with total space

$$f^*E = \{(x, l) \in M \times E | f(x) = p(l)\},$$

a closed submanifold of $M \times E$, and projection $f^*p : f^*E \to M$ given by $(f^*p)(x, l) = x$.

Bibliographical remarks. *We recommend [BT] to get some intuition for vector bundles and [AMR] for many technical details.*

2.4 The tangent bundle

Definition. Let M be a manifold and p in M.

(1) A "curve at p" is a smooth map $\gamma : I \to M$, where I is an open intervall in $\mathbb{R}$ containing 0, such that $\gamma(0) = p$.

(2) Two curves γ_1 and γ_2 at p are called "tangent with respect to the chart (U, φ)" if p is in U and $\frac{d}{dt}\Big|_0 (\varphi \circ \gamma_1)(t) = \frac{d}{dt}\Big|_0 (\varphi \circ \gamma_2)(t)$.

Lemma. *Let $(U, \varphi), (U', \varphi')$ be charts on a manifold M and p in $U \cap U'$. Then two curves γ_1 and γ_2 are tangent with respect to (U, φ) if and only if they are tangent with respect to (U', φ').*

Proof. Let γ_1 and γ_2 be equivalent with respect to (U, φ). Then

$$\frac{d}{dt}\Big|_0 (\varphi' \circ \gamma_1)(t) = \frac{d}{dt}\Big|_0 ((\varphi \circ \varphi^{-1}) \circ (\varphi \circ \gamma_1))(t) = (D_{\varphi(p)}(\varphi' \circ \varphi^{-1}))\frac{d}{dt}\Big|_0 (\varphi \circ \gamma_1)(t)$$

$$= \Big(D_{\varphi(p)}(\varphi' \circ \varphi^{-1})\Big)\frac{d}{dt}\Big|_0 (\varphi \circ \gamma_2)(t) = \frac{d}{dt}\Big|_0 (\varphi' \circ \gamma_2)(t).$$

Exchanging the roles of φ and φ' we arrive at the assertion. □

Remark. Let V be open in $\mathbb{R}^m$, V' open in $\mathbb{R}^n$ and $\Psi : V \to V'$ a smooth map. We recall that the "derivative of Ψ" in a point x in V is given by the unique linear map $D_x\Psi : \mathbb{R}^m \to \mathbb{R}^n$ such that for $w \in \mathbb{R}^m \backslash \{0\}$

$$\lim_{\|w\| \to 0} \frac{\|\Psi(x + w) - \Psi(x) - (D_x\Psi)(w)\|}{\|w\|} = 0.$$

Definitions. Let M be a manifold and let p be a point in M.

(1) Two curves γ_1 and γ_2 at p are called "equivalent" if γ_1 and γ_2 are tangent with respect to one (and thus every) chart.

(2) The space $\{[\gamma]_p \mid \gamma \text{ curve at } p\}$ of equivalence classes of curves at p is called the "tangent space to M at p" and is denoted by T_pM.

(3) The disjoint union $\dot{\bigcup}_{p \in M} T_pM$ is called the "tangent bundle of M" and is denoted by TM.

(4) Let N be a manifold and $f : M \to N$ a smooth map. The map $T_pf : T_pM \to T_{f(p)}N, [\gamma]_p \mapsto [f \circ \gamma]_{f(p)}$ is called the "tangent of f at p".

Remarks. (1) The tangent of a map f at p is well–defined. Proof as an exercise.)

(2) Despite the terminology it is not obvious that the tangent bundle is a vector bundle! (The proof will be given in the course of this section.)

(3) Given a smooth map $f : M \to N$ we have set–theoretically the "tangent of f"

$$Tf : TM \to TN, (Tf)([\gamma]_p) = (T_pf)([\gamma]_p).$$

Denoting the projection $TM \to M, [\gamma]_p \mapsto p$ by p_{TM} (and analogously for N) we obviously have $p_{TN} \circ Tf = f \circ p_{TM}$.

Lemma. *Let L, M, N be manifolds and $f : M \to N$ and $g : L \to M$ be smooth maps.*

(i) Then $T(f \circ g) = (Tf) \circ (Tg)$.

(ii) If $M = N$ and $f = \mathrm{Id}_M$ then $Tf = T\mathrm{Id}_M = \mathrm{Id}_{TM}$.

(iii) If f is a diffeomorphism then Tf is bijective and $(Tf)^{-1} = T(f^{-1})$.

Proof. Exercise. $\square$

Lemma. *Let V be open in $\mathbb{R}^m, x$ in V and w in $\mathbb{R}^m$. Let us set $\gamma_w(t) = x + tw$. Then*

(i) $[\gamma_{w'}]_x = [\gamma_{w''}]_x$ if and only if $w' = w''$ in $\mathbb{R}^m$, and

(ii) $\chi_x : \mathbb{R}^m \to T_xV, w \mapsto [\gamma_w]_x$ is bijective.

Remark. To be completely unambiguous one should note the curve $t \mapsto x + tw$ by $\gamma_{w,x}$. In order to simplify the notation we will stick to γ_w especially if we consider the equivalence classes $[\gamma_w]_x$, where no danger of confusion can arise.

Proof of the lemma.

Ad(i). Let us assume that $[\gamma_{w'}]_x = [\gamma_{w''}]_x$. Then $0 = \frac{d}{dt}\Big|_0(\gamma_{w'}(t) - \gamma_{w''}(t)) = w' - w''$. Thus the first assertion is proven.

Ad(ii). By (i) χ_x is injective. Let now γ be any curve at x and $w = \dot{\gamma}(0)$. It follows that $\frac{d}{dt}\Big|_0(\gamma(t) - \gamma_w(t)) = 0$, i.e. $[\gamma]_x = [\gamma_w]_x$. $\square$

Proposition. *Let V be open in $\mathbb{R}^m$. Then*

(i) T_xV is a vector space with a basis given by $[\gamma_{e_k}]_x$ for $k = 1, \ldots, m$. (The vector e_k is of course the k–th unit vector in R^m.)

(ii) TV carries naturally the structure of a trivial vector bundle over V via the bijection

$$\chi_V = V \times \mathbb{R}^m \to TV, \chi_V(x, w) = \chi_x(w) = [\gamma_w]_x,$$

fulfilling $p_{TV} \circ \chi_V = pr_1$.

(iii) If V' is open in $\mathbb{R}^n$ and $f : V \to V'$ is a smooth map then $Tf : TV \to TV'$ is a smooth vector bundle homomorphism over f. Identifying TV with $V \times \mathbb{R}^m$ and TV' with $V' \times \mathbb{R}^n$ we have $(Tf)(x, w) = (T_xf)(w) = (x, (D_xf)(w))$.

(iv) If a map f as in (iii) is a diffeomorphism then Tf is a vector bundle isomorphism.

Proof. The first assertion follows immediately from the preceding lemma.

Let $\chi_V : V \times \mathbb{R}^m \to TV$ be given by $\chi_V(x, w) = [\gamma_w]_x$ then we have $\chi_V \circ p_{TV} = \mathrm{Id}_V \circ pr_1$ and thus TV carries canonically the structure of a smooth real vector bundle of rank m, proving (ii).

Since $(Tf)[\gamma_w]_x = [f \circ \gamma_w]_{f(x)}$ and $\frac{d}{dt}\Big|_0(f \circ \gamma_w)(0) = (D_xf)(w)$ we arrive at

$$(Tf)[\gamma_w]_x = [\gamma_{(D_xf)(w)}]_{f(x)}.$$

This implies with $\Psi_V = (\chi_V)^{-1}$ and $\Psi'_{V'} = (\chi_{V'})^{-1}$ that

$$\Psi'_{V'} \circ Tf \circ (\Psi_V)^{-1} : V \times \mathbb{R}^m \to V' \times \mathbb{R}^n$$

is the map $(x, w) \mapsto (f(x), (D_x f)(w))$. Obviously this map is a smooth vector bundle homomorphism over the map $f : V \to V'$.

The last assertion follows from (iii) and the properties we have shown for Tf as a set–theoretic map. $\square$

Remark. Let V be open in $\mathbb{R}^m, V'$ open in $\mathbb{R}^n$ and $f : V \to V'$ smooth. Denoting the canonical basis of $\mathbb{R}^m$ by $\{e_1, \ldots, e_m\}$ and the canonical basis of $\mathbb{R}^n$ by $\{\epsilon_1, \ldots, \epsilon_n\}$ we have for each x in V a matrix A_x in $\mathrm{Mat}(n \times m, \mathbb{R})$ such that $(T_x f)(e_j) = \sum_{i=1}^{n} (A_x)_{ij} \epsilon_i$. Writing $f = {}^t(f_1, \ldots, f_n)$ with n scalar functions f_i it follows that $(A_x)_{ij} = \frac{\partial f_i}{\partial x_j}(x)$, i.e. A_x is the Jacobi matrix of f in x. (Fill in the details of the computation as an exercise.)

Definition. Let M be a manifold and $\varphi : U \to V \subset \mathbb{R}^m$ a chart. We call the bijection $T\varphi : TU \to TV$ the "natural bundle chart (associated to the chart (U, φ))".

Proposition. *Let M be a manifold and $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) | \alpha \in A\}$ an admissible atlas for M. Then the "natural bundle atlas associated to $\mathfrak{A}$", $T\mathfrak{A} = \{(TU_\alpha, T\varphi_\alpha) | \alpha \in A\}$ is a smooth atlas on TM such that TM together with its canonical projection p_{TM} to M is a smooth vector bundle over M.*

Furthermore the differentiable structure of $(TM, T\mathfrak{A})$ depends only on the differentiable structure of $(M, \mathfrak{A})$.

Proof. Since the TU_α cover set–theoretically TM and $T\varphi_\alpha : TU_\alpha \to TV_\alpha$ ($V_\alpha = \varphi(U_\alpha) \subset \mathbb{R}^m$) is bijective for each α we can define a topology $\mathcal{T}$ on TM as usual, i.e. Ω is open in TM if and only if $(T\varphi_\alpha)(\Omega \cap TU_\alpha)$ is open in TV_α for all α in A. The continuity of p_{TV_α} implies that $p_{TM} : TM \to M$ is continuous, and it follows that $(TM, \mathcal{T})$ is Hausdorff and second–countable.

The coordinate changes of the atlas $T\mathfrak{A}$ are given by

$$(T\varphi_\beta) \circ (T\varphi_\alpha)^{-1} = T(\varphi_\beta \circ \varphi_\alpha^{-1}) : (T\varphi_\alpha)(TU_{\alpha\beta}) \to (T\varphi_\beta)(TU_{\alpha\beta}).$$

Since $T(\varphi_\beta \circ \varphi_\alpha^{-1})$ is a diffeomorphism, it follows that the atlas $T\mathfrak{A}$ on TM is smooth, and furthermore that p_{TM} is a smooth map. Since TV_α is isomorphic as a vector bundle over V_α to the trivial vector bundle $V_\alpha \times \mathbb{R}^m$ via the vector bundle isomorphism $(\chi_\alpha)^{-1} := (\chi_{V_\alpha})^{-1}$, for a given α, TU_α inherits the structure of a vector bundle from TV_α. (We denote the map $V_\alpha \times \mathbb{R}^m \to TV_\alpha, (x, w) \mapsto [\gamma_w]_x$ again by χ_{V_α}.) Using the fact that the transition maps $T(\varphi_\beta \circ \varphi_\alpha^{-1})$ are vector bundle isomorphisms, the linear structures on the fibers of $p_{TM} : TM \to M$ are well–defined and

$$(\varphi_\alpha^{-1} \circ pr_1, pr_2) \circ \chi_\alpha \circ T\varphi_\alpha : TU_\alpha \to U_\alpha \times \mathbb{R}^m$$

are local trivializiations of TM over U_α. $\square$

Exercise. Show the last assertion of the proposition, i.e. the independence of this construction of the chosen admissible atlas given a fixed differentiable structure on M.

Proposition. *Let M and N be manifolds and $f : M \to N$ a smooth map. Then*

(i) $Tf : TM \to TN$ is a smooth vector bundle homomorphism over f, and

(ii) if $\varphi_\alpha = (x_1^\alpha, \ldots, x_m^\alpha) : U_\alpha \to V_\alpha \subset \mathbb{R}^m$ respectively $\psi_\beta = (y_1^\beta, \ldots, y_n^\beta) : U'_\beta \to V'_\beta \subset \mathbb{R}^n$ are local coordinates near p in M respectively $f(p)$ in N, then the map $T_p f : T_p M \to T_{f(p)} N$ is given by the "Jacobi–matrix of f in the local coordinate systems", i.e. setting

$$^t(\tilde{f}_1, \ldots, \tilde{f}_n) = \tilde{f} = \psi_\beta \circ f \circ \varphi_\alpha^{-1} : \ \varphi_\alpha(U_\alpha \cap f^{-1}(U'_\beta)) \to \mathbb{R}^n$$

and denoting the canonical basis of $\mathbb{R}^m$ and $\mathbb{R}^n$ again by $\{e_1, \ldots, e_m\}$ and $\{\epsilon_1, \ldots, \epsilon_n\}$ we have

$$T\psi_\beta \circ Tf \circ T\varphi_\alpha^{-1} = T\tilde{f} \quad \textit{and}$$

$$T\tilde{f}([\gamma_{e_j}]_{\varphi_\alpha(p)}) = \sum_{i=1}^{n} \frac{\partial \tilde{f}_i}{\partial x_j^\alpha}(\varphi_\alpha(p)) \cdot [\mu_{\epsilon_i}]_{\Psi_\beta(f(p))},$$

where, as usual, $[\gamma_{e_j}]_x$ is the tangent vector in x represented by $\gamma_{e_j}(t) = x + t_{e_j}$ and analogously for $[\mu_{\epsilon_i}]_y$.

Proof. Exercise. □

Remark. Obviously the basic theorems of differential calculus in several variables can now be applied to smooth manifolds and yield for example the following

Proposition. *Let $f : M \to N$ be a smooth map and q in N. If the rank of $T_p f$ is either maximal for all p in $f^{-1}(q)$ or constant on a neighborhood of $f^{-1}(q)$ in M then $f^{-1}(q)$ is a (smooth) closed submanifold of M.*

Proof. Exercise using the rank theorem in local coordinates. □

Bibliographical remarks. *Beside the books mentioned at the end of Section 2.1 we would like to recommend [BröJä] and [Brö] - the latter being unfortunately available only in german language.*

2.5 Vector fields on manifolds

Definitions. Let M be a manifold.

(1) The vector space of real–valued smooth functions on M is denoted by $\mathcal{E}(M) = C^\infty(M, \mathbb{R})$.

(2) Let U_f and U_g be open neighborhoods of p in M and $f \in \mathcal{E}(U_f), g \in \mathcal{E}(U_g)$. We say "$f$ and g are equivalent in p", $f \sim_p g$, if and only if there is an open neighborhood V of p such that $V \subset U_f \cap U_g$ and $f|_V = g|_V$. The set of equivalence classes

$$\{f : U_f \to \mathbb{R} | p \in U_f, U_f \text{ open in } M, f \text{ smooth }\}/_{\sim_p}$$

is denoted by $\mathcal{E}_p(M)$ and an equivalence class $[f : U_f \to \mathbb{R}]_p$ is denoted by $\underset{\sim p}{f}$ and called a "germ of a smooth function in p".

Proposition. *Let M be a manifold and p in M. Then*

(i) $\mathcal{E}_p(M)$ is a commutative, associative, unital $\mathbb{R}$–algebra, and

(ii) the vector space $Der(\mathcal{E}_p(M)) :=$

$\{D_p : \mathcal{E}_p(M) \to \mathbb{R} | D_p$ *is* $\mathbb{R}$*–linear and* $D_p(\underset{\sim p}{f} \cdot \underset{\sim p}{g}) = D_p(\underset{\sim p}{f}) \cdot g(p) + f(p) \cdot D_p(\underset{\sim p}{g})\}$

of its "(scalar–valued) derivations" is $\mathbb{R}$–linearly isomorphic to T_pM.

Proof. The first assertion follows from the observation that $\mathcal{E}(U)$ is a commutative, associative, unital $\mathbb{R}$–algebra for each open set U in M.

Since $\mathcal{E}(U_\alpha) \to \mathcal{E}(V_\alpha), f \mapsto f \circ \varphi_\alpha^{-1}$ is a $\mathbb{R}$–algebra isomorphism for each chart $\varphi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^m$, we can assume that $M = \mathbb{R}^m$ and $p = 0$.

The map

$$T_0\mathbb{R}^m \xrightarrow{\delta_0} \text{Der}\,(\mathcal{E}_0(\mathbb{R}^m)),\ \delta_0([\gamma]_0)(\underset{\sim 0}{f}) = \frac{d}{dt}\Big|_0 (f \circ \gamma)(t)$$

is well–defined and an injective $\mathbb{R}$–vector space homomorphism. In order to show that δ_0 is surjective we use the "Fundamental lemma" below to develop a smooth function f near 0 in $\mathbb{R}^m$ as follows

$$f(x) = f(0) + \sum_{j=1}^m f_j(x) \cdot x_j,$$

where f_j is a smooth function near 0 fulfilling $f_j(0) = \frac{\partial f}{\partial x_j}(0)$. Given now D_0 in Der$\,(\mathcal{E}_0(\mathbb{R}^m))$ we set $a_j = D_0(x_j) \in \mathbb{R}$, $w = \sum_{j=1}^m a_j e_j$, and $\gamma_w(t) = t \cdot w$ as usual. Then we have

$$\delta_0([\gamma_w]_0)(\underset{\sim 0}{f}) = \frac{d}{dt}\Big|_0 f(tw) = \sum_{j=1}^m a_j \cdot \frac{\partial f}{\partial x_j}(0) = \sum_{j=1}^m f_j(0) D_0(x_j) = D_0(\underset{\sim 0}{f}). \qquad \square$$

Lemma ("Fundamental lemma").

Let U be an open neighborhood of 0 in $\mathbb{R}^m$ and $f : U \to \mathbb{R}$ a smooth function. Then there are smooth function f_j defined near 0 (for $j = 1, \ldots, m$) such that $f_j(0) = \frac{\partial f}{\partial x_j}(0)$ and

$$f(x) = f(0) + \sum_{j=1}^m f_j(x) \cdot x_j$$

for all x in a ball $\mathbb{B}_\epsilon(0)$ for $\epsilon > 0$ sufficiently small.

Proof.

Let $\epsilon > 0$ such that the ball $\mathbb{B}_\epsilon(0)$ is contained in U. For all x in $\mathbb{B}_\epsilon(0)$ we have

$$f(x) - f(0) = \int_0^1 \frac{d}{dt} f(t \cdot x) dt = \sum_{j=1}^m \Big(\int_0^1 \frac{\partial f}{\partial x_j}(t \cdot x) dt \Big) \cdot x_j.$$

Obviously $f_j(x) := \int_0^1 \frac{\partial f}{\partial x_j}(t \cdot x)dt$ is a smooth function on $\mathbb{B}_\epsilon(0)$ and $f_j(0) = \frac{\partial f}{\partial x_j}(0)$. $\square$

Corollary. *Let V be open in $\mathbb{R}^m$ and x in V. Then the isomorphism*

$$T_xV \xrightarrow{\delta_x} Der(\mathcal{E}_x(V)),\ \delta_x([\gamma]_x)(\underset{\sim x}{f}) = \frac{d}{dt}\Big|_0 (f \circ \gamma)(t)$$

maps $[\gamma_{e_k}]_x$ to the partial derivative $\frac{\partial}{\partial x_k}\Big|_x$.

Proof. Exercise. $\square$

Definition. Let M be a manifold. A "vector field (on M)" is a smooth section of the tangent bundle TM of M. We denote the $\mathbb{R}$–vector space of all vector fields by $\mathfrak{X}(M) = \Gamma_{C^\infty}(M, TM)$.

Proposition. *Let V be open in $\mathbb{R}^m$ and let*

$$\mathrm{Der}(\mathcal{E}(V)) = \{D : \mathcal{E}(V) \to \mathcal{E}(V) | D \text{ is } \mathbb{R}\text{–linear and} \quad D(f \cdot g) = D(f) \cdot g + f \cdot D(g)\}$$

be the space of all "derivations of $\mathcal{E}(V)$". Then the maps $\delta_x : T_xV \to Der_x(\mathcal{E}_x(V))$ yield a $\mathbb{R}$–vector space isomorphism $\delta : \mathfrak{X}(V) \to Der(\mathcal{E}(V))$.

Proof. Identifying TV via χ_V with $V \times \mathbb{R}^m$, a vector field X on V is given by a smooth map $w : V \to \mathbb{R}^m$ ($X(x) = [\gamma_{w(x)}]_x$ in T_xV). The map δ_x takes $X(x)$ to $\sum_{j=1}^m w_j(x) \frac{\partial}{\partial x_j}\Big|_x$, where $w = {}^t(w_1, \ldots, w_m)$. Obviously $D^X := \sum_{j=1}^m w_j \frac{\partial}{\partial x_j}$ is an element of $\mathrm{Der}(\mathcal{E}(V))$, since the functions w_j are smooth, and furthermore we observe that the map $\delta : \mathfrak{X}(V) \to \mathrm{Der}(\mathcal{E}(V))$, $\delta(X) = D^X$ is injective. Let now D be in $\mathrm{Der}(\mathcal{E}(V))$, then applying D to germs of smooth functions yields a map $D_x : \mathcal{E}_x(V) \to \mathbb{R}, D_x(\underset{\sim x}{f}) = (D(f))(x)$, which is an element of $\mathrm{Der}_x(\mathcal{E}_x(V))$. Let now the functions w_j be defined by $w_j(x) = D_x(x_j)$ (x_j being again the j–th coordinate on $V \subset \mathbb{R}^m$). Since $D_x(x_j) = D(x_j)(x)$, these functions are smooth and thus define a vector field X on V. It follows that for f in $\mathcal{E}(V)$

$$D^X(f)(x) = D_x^X(\underset{\sim x}{f}) = \sum_{j=1}^m \Big(\frac{\partial f}{\partial x_j}\Big|_x\Big) w_j(x) = D_x(\underset{\sim x}{f})$$

by applying the fundamental lemma to functions defined near x. Since D is uniquely determined by the family $\{D_x | x \in V\}$ it follows that $D = D^X = \delta(x)$ i.e. δ is an isomorphism.

$\square$

Corollary. *Let M be a manifold. Then there is a $\mathbb{R}$–vector space isomorphism $\delta : \mathfrak{X}(M) \to Der(\mathcal{E}(M))$ such that in all charts (V, φ) of M δ is given as in the preceding proposition.*

Proof. Since a vector field as well as a derivation of $\mathcal{E}(M)$ are uniquely determined by their restrictions to the chart domains in an atlas of M, the map δ is well–defined and injective. Given now D in $\mathrm{Der}\,(\mathcal{E}(M))$ we construct X_α on U_α for an atlas $\{(U_\alpha, \varphi_\alpha) | \alpha \in A\}$ as in the proof of the preceding lemma. It remains only to show that the X_α define a global vector field, i.e. a section of TM. This easily follows from the transformation properties of the partial derivatives defined on coordinate charts and the definition of the bundle charts $T\varphi_\alpha$ of TM. (Details as exercise.) $\square$

Remark. The last proposition explains why one speaks of a "vector field on M, locally given by $\sum_{j=1}^n w_j^\alpha \frac{\partial}{\partial x_j^\alpha}$", where $\varphi_\alpha = (x_1^\alpha, \ldots, x_m^\alpha)$ are local coordinates and $w_j^\alpha : V_\alpha = \varphi_\alpha(U_\alpha) \to \mathbb{R}$ are smooth functions.

Definition. Let $\mathbb{K}$ be a field and $\mathfrak{g}$ be a $\mathbb{K}$–vector space with a map $[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. We call $(\mathfrak{g}, [,])$ a "($\mathbb{K}$–)Lie algebra" if the following conditions are satisfied:

(1) $[,]$ is $\mathbb{K}$–bilinear,

(2) $[,]$ is anti–symmetric,

(3) $[,]$ fulfills the "Jacobi identity":

$$[u, [v, w]] = [[u, v], w] + [v, [u, w]]$$

for all u, v, w in $\mathfrak{g}$.

Example (and exercise). Let $\mathcal{A}$ be an associative $\mathbb{K}$–algebra. Then $[S, T] := S \cdot T - T \cdot S$ defines a $\mathbb{K}$–Lie algebra structure on $\mathcal{A}$. Thus for a $\mathbb{K}$–vector space E the algebra $\mathcal{A} = \mathrm{End}_{\mathbb{K}}(E)$ is a $\mathbb{K}$–Lie algebra.

Remark. Originating in the above example, the name "commutator" for the map $[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is frequently used, even if $\mathfrak{g}$ ist *not* constructed from an associative algebra.

Definition. Let $(\mathfrak{g}, [,])$ be a $\mathbb{K}$–Lie algebra and $\mathfrak{h} \subset \mathfrak{g}$ a subset. We call $\mathfrak{h}$ a "Lie subalgebra (of $\mathfrak{g}$ over $\mathbb{K}$)" if $\mathfrak{h}$ is a $\mathbb{K}$–subspace of $\mathfrak{g}$ such that $[\xi, \eta]$ is in $\mathfrak{h}$ for all ξ, η in $\mathfrak{h}$.

Lemma. *Let A be a $\mathbb{K}$–algebra and $Der(A) := \{D \in End_{\mathbb{K}}(A) | D(a \cdot b) = D(a) \cdot b + a \cdot D(b)$ $\forall\, a, b \in A\}$. Then $Der(A)$ is a Lie subalgebra of $(End_{\mathbb{K}}(A), [,])$, the space of $\mathbb{K}$–linear vector space endomorphisms of A with the commutator $[,]$ coming from the associative composition of endomorphisms (as in the preceding example).*

Proof. Direct calculation (exercise). $\square$

Corollary 1. *Let M be a manifold. Then the space $Der(\mathcal{E}(M))$ of the associative $\mathbb{R}$–algebra $\mathcal{E}(M)$ is a Lie subalgebra of $(End_{\mathbb{R}}(\mathcal{E}(M)), [,])$.*

Proof. Follows directly from the preceding lemma. $\square$

Corollary 2. *Let M be a manifold. Then the $\mathbb{R}$–vector space $\mathfrak{X}(M)$ naturally carries the structure of a Lie algebra induced from $Der(\mathcal{E}(M))$.*

Proof. Let X, Y be in $\mathfrak{X}(M)$ and $D = [D^X, D^Y]$ in $\mathrm{Der}(\mathcal{E}(M))$. Since $\delta : \mathfrak{X}(M) \to \mathrm{Der}(\mathcal{E}(M))$ is an isomorphism there is a unique Z in $\mathfrak{X}(M)$ such that $D^Z = \delta(Z) = [D^X, D^Y]$. The bilinear map

$$[X, Y] := \delta^{-1}([\delta(X), \delta(Y)]).$$

obviously defines a Lie bracket. $\square$

Definitions. Let M be a manifold and X in $\mathfrak{X}(M)$.

(1) Let I be a connected open neighborhood of 0 in $\mathbb{R}$. A smooth curve $\gamma : I \to M$ is called a "(local) integral curve of X with initial condition p (in M)" if the following conditions are satisfied

$$\dot\gamma(t) = \frac{d}{dt}\gamma(t) = (T_t\gamma)\Big(\frac{d}{dt}\Big|_t\Big) = X(\gamma(t)) \quad \forall t \in I \quad \text{and} \quad \gamma(0) = p.$$

(2) Let Ω be an open set in $\mathbb{R} \times M$ containing $\{0\} \times M$. A smooth map $\varphi^X : \Omega \to M$ is called a "local flow of the vector field X" if for each p in M the curve $t \mapsto \varphi^X(t,p) =: \varphi^X_t(p)$ is an integral curve of X with initial condition p.

Theorem. *Let M be a manifold and X a vector field on M. Then for each p in M there exist local integral curves of X with initial condition p. Furthermore these curves are unique, i.e. they coincide on the intersection of their intervals of definition, and there is a maximal connected open interval I_p in $\mathbb{R}$ containing 0 such that $\varphi^X_t(p)$ is defined on I_p and cannot be extended beyond I_p.*

The set $\Omega := \{(t,p) \in \mathbb{R}\times M | t \in I_p\}$ is open in $\mathbb{R}\times M$ and the map $\varphi^X : \Omega \to M, \varphi^X(t,p) = \varphi^X_t(p)$ is smooth, i.e. there is a unique maximal local flow of the vector field.

Proof. Given any chart $\varphi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^m$ the vector field X corresponds to a map $w^\alpha : V_\alpha \to \mathbb{R}^m$ such that $X(x) = (x, w^\alpha(x))$ on V_α. The equation $\dot\gamma = X(\gamma(t))$ then translates to the ordinary differential equation:

$$\dot\gamma_k(t) = w^\alpha_k(\gamma(t)) \quad \text{for} \quad k = 1, \ldots, m.$$

Since the coefficient functions w^α_k are smooth there is for any point of V_α an open neighborhood and some $\epsilon > 0$ such that the solution of this differential equation exists, is smooth and unique and depends smoothly on the initial condition in this neighborhood. This local flow can then be transported to M and by uniqueness of the solutions they patch smoothly with solutions coming from other charts. Thus for each p in M there is a maximal connected open intervall I_p on which the curve $t \mapsto \varphi^X_t(p)$ is defined. Openness of the set Ω follows again from local arguments in charts. $\square$

Exercise. Fill out the details of the preceding proof (by using any textbook on ordinary differential equations for example.)

Corollary 1. ("Flow equations"). *Let $(M, X, \Omega, \varphi^X)$ be as in the theorem. Then $\varphi^X_0 = Id_M$, and if $(s,p), (t+s,p), (t, \varphi^X_s(p))$ are in Ω then $\varphi^X_{t+s}(p) = \varphi^X_t(\varphi^X_s(p))$, in short*

$$\text{“}\varphi^X_0 = Id_M \quad \textit{and} \quad \varphi^X_{t+s} = \varphi^X_t \circ \varphi^X_s.\text{”}$$

Proof. The curves $t \mapsto \varphi^X_{t+s}(p)$ and $t \mapsto \varphi^X_t(\varphi^X_s(p))$ are both defined for small t and are integral curves of X with initial condition $\varphi^X_s(p)$. Uniqueness of this integral curve yields the assertion. $\square$

Corollary 2. *Let $(M, X, \Omega, \varphi^X)$ be as in the theorem, t in $\mathbb{R}$ and U be open in M such that $\{t\} \times U$ is in Ω. Then $\varphi^X_t : U \to M$ is a diffeomorphism onto $\varphi^X_t(U)$.*

Proof. Obviously $t \in I_p$ for all p in U. By Corollary 1 we have for p in U and s in $[0,t]$ the equality $\varphi^X_{t-s}(p) = \varphi^X_{-s}(\varphi^X_t(p))$, i.e. φ^X_{-t} is defined on $\varphi^X_t(U)$ and $\varphi^X_{-t} = (\varphi^X_t|_U)^{-1} : \varphi^X_t(U) \to U$. □

Definition. Let M be a manifold and X in $\mathfrak{X}(M)$. The "support of X" is the closed subset of M defined by

$$\text{supp }(X) = \overline{\{p \in M | X(p) \neq 0\}}.$$

Corollary 3. *Let M be a manifold and X a vector field on M with compact support. Then $\Omega = \mathbb{R} \times M$, i.e. the "flow of X is global". The map $\varphi^X : \mathbb{R} \times M \to \mathbb{R}$ is then a smooth action of the Lie group $(\mathbb{R}, +)$ on M.*

Proof. Let p be in $M \backslash \text{supp }(X)$ then $\varphi^X_t(p) = p$, i.e. the unique maximal integral curve is defined for all t in $\mathbb{R}$. Since Ω is open in M there is for each p in M some $\epsilon_p > 0$ and an open neighborhood U_p of p in M such that $(-\epsilon_p, \epsilon_p) \times U_p$ is in Ω. Covering the compact set $K = \text{supp }(X)$ by $\{K \cap U_p | p \in K\}$ we find $\epsilon_0 > 0$ such that $(-\epsilon_0, \epsilon_0) \times M$ is in Ω. Let now $\epsilon_{\max} \geq \epsilon_0 > 0$ be the supremum of all $\epsilon > 0$ such that $(-\epsilon, \epsilon) \times M \subset \Omega$. Assuming that ϵ is not $+\infty$ there is a p in M such that I_p is at least unilaterally bounded. Without loss of generality we may assume that $I_p \subset (-\infty, 3 \cdot \epsilon_{\max}/2)$. Setting $\gamma(t) = \varphi^X_{t/2}(\varphi^X_{t/2}(p))$ for t in $[0, 2\epsilon_{\max})$ we see that the assumption on I_p is wrong and therefore $\epsilon_{\max} = +\infty$, i.e. the flow is global.

The two preceding corollaries now imply that $\varphi^X : \mathbb{R} \times M \to M$ is a smooth action. □

Remarks. (1) The third corollary shows in particular that the flow of a vector field on a compact manifold is always globally defined.

(2) The flow equations (see Corollary 1) are the appropriate formulation for a "local $\mathbb{R}$–action".

Lemma. *Let M be a manifold and let A be an open subset in $\mathbb{R} \times M$ such that $A \cap (\mathbb{R} \times \{p\})$ is a connected open interval containing 0 in $\mathbb{R} \times \{p\}$ for all p in M. Let furthermore $\varphi : A \to M$ be a smooth map such that the flow equations of Corollary 3 are satisfied. Then the vectorfield X defined by $X(p) := \frac{d}{dt}\big|_0 \varphi^X(t,p)$ is smooth on M and the maximal flow φ^X is defined on an open set $\Omega \subset \mathbb{R} \times M$ such that $A \subset \Omega$ and $\varphi^X|_A = \varphi$.*

Proof. Exercise. □

Remarks. A vector field X on a manifold M together with its flow φ^X is often called a "(continuous) dynamical system on M". Rigorously speaking this term should be reserved for those vector fields whose flow is globally (i.e. on $\mathbb{R} \times M$) defined. A "discrete dynamical system" is a $\mathbb{Z}$–action $\vartheta : \mathbb{Z} \times M \to M$. Since $\vartheta_0 = \text{Id}_M, \vartheta_{-1} = (\vartheta_1)^{-1}$ and $\vartheta_n = (\vartheta_1)^n$ the action is in fact determined by the diffeomorphism $\vartheta_1 : M \to M$. Allowing any smooth map $f : M \to M$ to replace ϑ_1 we arrive at the notion of a "semi–group action of $\mathbb{N}_0$ on M":

$$\vartheta : \mathbb{N}_0 \times M \to M, \vartheta(n,p) = (\underbrace{\varphi \circ \cdots \circ \varphi}_{n\,\text{factors}})(p) = \varphi^n(p)$$

(and we set of course $\vartheta(0,p) = p$, i.e. $\varphi^0 := \vartheta_0 := \text{Id}_M$).

In closing this section let us make contact to the Hamilton equation in $\mathbb{R}^{2n}$.

Proposition. *Let M be open in $\mathbb{R}^{2n}$ with coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$. Then the Hamilton equation associated to a smooth function $H : M \to \mathbb{R}$, given as*

$$\dot{q}_j = \frac{\partial H}{\partial p_j}, \quad \dot{p}_j = -\frac{\partial H}{\partial q_j} \quad \textit{with initial condition} \quad (q^0, p^0) \quad \textit{in} \quad M$$

are equivalent to

$$\dot{\gamma}(t) = X_H(\gamma(t)) \quad \textit{and} \quad \gamma(0) = (q^0, p^0)$$

for the vector field $X_H(q,p) = \sum_{j=1}^n \left(\frac{\partial H}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial H}{\partial q_j} \frac{\partial}{\partial p_j} \right)$.

Proof. Exercise. □

Remark. The vector field X_H in the last proposition is called the "Hamiltonian vector field" associated to the Hamilton function H.

Bibliographical remarks. *As at the end of the previous section, plus a solid reference on ordinary differential equations as [Ar2].*

2.6 Differential forms and the Lie derivative

Definitions. Let M be a manifold and TM its tangent bundle.

(1) The vector bundle $(TM)^* =: T^*M \stackrel{p_{T^*M}}{\longrightarrow} M$ is called the "cotangent bundle of M" and the vector space $(p_{T^*M})^{-1}(x) = (T_xM)^* =: T_x^*M$ the "cotangent space in x" ($x \in M$).

(2) A section of T^*M is called a "(differential) one–form on M" and the vector space of all its sections is denoted by $\mathcal{E}^1(M) = \Gamma_{C^\infty}(M, T^*M)$.

(3) Analogously we define for $k \geq 1$ the bundles $\Lambda^k T^*M := \Lambda^k (TM)^*$ and call their sections "(differential) k–forms on M". The section spaces are denoted by $\mathcal{E}^k(M) := \Gamma_{C^\infty}(M, \Lambda^k T^*M)$.

(4) The space of smooth functions $\mathcal{E}(M) = C^\infty(M, \mathbb{R})$ is also called the space of "0–forms", $\mathcal{E}^0(M) := \mathcal{E}(M)$.

Remark. If $m = \dim_{\mathbb{R}} M$, then $\Lambda^k T_x^* M = \{0\}$ for all x in M and all $k > m$ and thus there are no (non–trivial) k–forms with $k > \dim_{\mathbb{R}} M$.

Definition. Let V be open in $\mathbb{R}^m$ with coordinates $(x_1, \ldots, x_m)$ and let a vector field X be identified with the associated derivation $\delta(X) = \sum_{j=1}^m a_j \frac{\partial}{\partial x_j}$ ($a_j \in \mathcal{E}(V)$). The differential one–form $dx_j \in \mathcal{E}^1(V)$ is defined by $(dx_j)_p(\frac{\partial}{\partial x_k}|_p) = \delta_{j,k}$ for all p in V and thus $dx_j(X) = a_j$.

Lemma. *Let V be open in $\mathbb{R}^m$ and p in V. Then*

(i) $T_p^*V = ((\{(dx_1)_p, \ldots, (dx_m)_p\}))_{\mathbb{R}}$, and

(ii) $\Lambda^k T_p^* V = ((\{dx_{i_1} \wedge \ldots \wedge dx_{i_k})_p | 1 \leq i_1 \leq \cdots \leq i_k \leq n\}))_{\mathbb{R}}$.

Furthermore the section space $\mathfrak{X}(\mathfrak{V}), \mathcal{E}^\circ(\mathfrak{V})$ *and* $\mathcal{E}^k(V)$ *for* $k \geq 1$ *are free* $\mathcal{E}(V)$*–modules with module basis as follows*

(iii) $\mathfrak{X}(V) = ((\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}\}))_{\mathcal{E}(V)}$,

(iv) $\mathcal{E}^\circ(V) = ((\{1\}))_{\mathcal{E}(V)}$, *and*

(v) $\mathcal{E}^k(V) = ((\{dx_{i_1} \wedge \cdots \wedge dx_{i_k} | 1 \leq i_1 \leq \cdots \leq i_k \leq n\}))_{\mathcal{E}(V)}$.

Proof. Exercise. □

Definition. Let M be a manifold, p in M, v in T_pM and η in $\Lambda^k T_p^* M$ ($k \geq 1$). Then the "contraction of v and η" is the $(k-1)$–form $v \lrcorner \eta = i_v \eta \in \Lambda^{k-1} T_p^* M$ defined as follows:

$$(i_v \eta)(v_1, \ldots, v_{k-1}) := \eta(v, v_1, \ldots, v_{k-1}) \quad \text{for all} \quad v_1, \ldots, v_{k-1} \quad \text{in} \quad T_pM.$$

Exercise. Check that $i_v \eta$ is multilinear and alternating, i.e. in $\Lambda^{k-1} T_p^* M$.

Lemma. *Let* M *be a manifold,* X *in* $\mathfrak{X}(M)$ *and* η *in* $\mathcal{E}^k(M)$*, then the contraction* $X \lrcorner \eta = i_X \eta$ *defined by*

$$(i_X \eta)_p := i_{X_p} \eta_p \quad \textit{for all} \quad p \quad \textit{in } M$$

is a smooth differential $(k-1)$*–form, i.e.* $i_X \eta$ *is in* $\mathcal{E}^{k-1}(M)$*.*

Proof. Exercise. □

Remark. On the space $\Lambda^*(T_p^* M) = \bigoplus_{k \geq 0} \Lambda^k(T_p^* M)$ one has the multiplication "$\wedge$" of exterior algebras. This is easily globalized as follows.

Lemma. *Let* M *be a manifold and* $\eta \in \mathcal{E}^k(M), \mu \in \mathcal{E}^l(M)$*. Then the formula*

$$(\eta \wedge \mu)_p := \eta_p \wedge \mu_p \quad \textit{for all} \quad p \textit{ in } M$$

defines a smooth $(k+l)$*–form* $\eta \wedge \mu$ *on* M*.*

Furthermore the space of sections of the vector bundle $\Lambda^* T^* M := \bigoplus_{k \geq 0} \Lambda^k T^* M$ *is canonically isomorphic to* $\bigoplus_{k \geq 0} \mathcal{E}^k(M)$ *and this space together with the wedge–product is a super–commutative, associative, unital algebra over the ring* $\mathcal{E}(M)$*.*

Proof. Exercise. □

Remark. Given a smooth map $f : M \to N$ between two manifolds M and N, and a point p in M we have the tangent of f in p, which we will denote also by $(f_*)_p$, i.e.

$$(f_*)_p := T_p f : T_p M \to T_{f(p)} N.$$

Since $T_p f$ is linear we have induced maps $\otimes^k T_{f(p)}^* N \to \otimes^k T_p^* M$ and $\Lambda^k T_{f(p)}^* N \to \Lambda^k T_p^* M$.

Definition. Let $f : M \to N$ be smooth and η in $\mathcal{E}^k(N)$. Then the k–form $f^* \eta$ is defined by

$$(f^* \eta)_p(v_1, \ldots, v_k) := \eta_{f(p)}((f_*)_p(v_1), \ldots, (f_*)_p(v_k)) \quad \forall\, p \in M, \ \forall\, v_1, \ldots, v_k \in T_p M$$

is called the "pull–back of η by f".

Exercise. Check that $f^*\eta$ is a smooth k–form on M.

Remark. It is important to note that the anologous construction on a vector field does *not* always work, i.e. $(f_*X)_p = (f_*)_p X_p \in T_{f(p)}N$ for p in M and X in $\mathfrak{X}(M)$, but this is in general only a section of the bundle $f^*(TN) \to M$ and does not necessarily define a vector field on N.

Example. Let $M = N = \mathbb{R}$ and $f(x) = x^2$. Then $(T_p f)(a(x)\frac{\partial}{\partial x}|_p) \in T_{f(p)}\mathbb{R}$ defines a vector field on $\mathbb{R}$ if and only if $a(x) = 0$ for all x in $\mathbb{R}$, i.e. if and only if the vector field $X = a(x)\frac{\partial}{\partial x}$ is everywhere zero.

Lemma. *Let L, M, N be manifolds and $g : L \to M$ and $f : M \to N$ be smooth. Furthermore let $\eta \in \mathcal{E}^k(N)$ and μ in $\mathcal{E}^l(N)$. Then*

(i) $f^*(\eta \wedge \mu) = (f^*\eta) \wedge (f^*\mu)$,

(ii) $f^*\psi = \psi \circ f$ for ψ in $\mathcal{E}^0(N)$,

(iii) $f^* : \mathcal{E}^*(N) \to \mathcal{E}^*(M)$ is an $\mathbb{R}$–linear even homomorphism of superalgebras (over $\mathbb{R}$) fulfilling

$$f^*(\psi \cdot \eta) = (f^*\psi) \cdot (f^*\eta) \quad \textit{for all} \quad \psi \textit{ in } \mathcal{E}^0(N) \ ,$$

(iv) $(f \circ g)^* = g^* \circ f^* : \mathcal{E}^*(N) \to \mathcal{E}^*(L)$.

Proof. Exercise. $\square$

Definition. Let M be a manifold, X a vector field on M and φ_t^X the flow of X.

For a k–form η on M we define the "Lie derivative of η with respect to X" by

$$(\mathcal{L}_X \eta)_p := \frac{d}{dt}\Big|_0 \Big(((\varphi_t^X)^*\eta)_p \Big) \quad \text{for all } p \text{ in } M.$$

Proposition. *Let M, X, φ_t^X be as in the preceding definition and let η, η' be in $\mathcal{E}^k(M)$ and μ in $\mathcal{E}^l(M)$ with $k, l \geq 1$. Then*

(i) $\mathcal{L}_X\eta$ is in $\mathcal{E}^k(M)$,

(ii) $\mathcal{L}_X(\lambda \cdot \eta) = \lambda \cdot \mathcal{L}_X\eta$ for λ in $\mathbb{R}$,

(iii) $\mathcal{L}_X(\eta + \eta') = \mathcal{L}_X\eta + \mathcal{L}_X\eta'$,

(iv) $\mathcal{L}_X(\eta \wedge \mu) = (\mathcal{L}_X\eta) \wedge \mu + \eta \wedge (\mathcal{L}_X\mu)$.

Proof.

Ad(i). Since the local flow φ^X is a smooth map from its domain of definition $\Omega \subset \mathbb{R} \times M$ to M it easy to deduce that $\mathcal{L}_X\eta$ exists in every point p of M and that $(\mathcal{L}_X\eta)_p$ is in $\Lambda^k(T_pM)^*$, i.e. $\mathcal{L}_X\eta$ is a section of $\Lambda^k T^*M$. Since differentiability of a section is a local condition we may assume that $M = V$ is an open subset of $\mathbb{R}^m$. The fact that $\mathcal{E}^k(V) = ((\{dx_{i_1} \wedge \cdots \wedge x_{i_k} | i_1 < \cdots < i_k\}))_{\mathcal{E}(V)}$ implies that $\mathcal{L}_X\eta$ is a smooth section if and only if $(\mathcal{L}_X\eta)(X_1,\ldots,X_k)$ is a smooth function for all $X_1,\ldots,X_k$ in $\mathfrak{X}(V)$. For p near a fixed point p_0 in V we have

$$((\mathcal{L}_X\eta)(X_1,\ldots,X_k))(p) = \frac{d}{dt}\Big|_0((\varphi_t^X)^*\eta)_p(X_1(p),\ldots,X_k(p))$$

$$= \frac{d}{dt}\Big|_0\Big(\eta_{\varphi_t^X(p)}(((\varphi_t^X)_*)_p(X_1(p)),\ldots((\varphi_t^X)_*)_p(X_k(p)))\Big) =: \frac{d}{dt}\Big|_0 F(t,p),$$

and for fixed $X, X_1,\ldots,X_k,\eta$ the function $F(t,p)$ is defined for small t and p near p_0 and is smooth in both variables. Thus $(\mathcal{L}_X\eta)(X_1,\ldots,X_k)$ is a smooth function in p near p_0 for all p_0 in V, and the first assertion is proven.

Assertions (ii) and (iii) follow directly from the $\mathbb{R}$–linearity of the maps $(\varphi_t^X)^*$.

The last assertion can be derived from the formula $(\varphi_t^X)^*(\eta\wedge\mu) = (\varphi_t^X)^*\eta\wedge(\varphi_t^X)^*\mu$ by writing $\frac{d}{dt}|_0((\varphi_t^X)^*\sigma)_p = \lim_{t\to 0}\frac{1}{t}(((\varphi_t^X)^*\sigma)_p - \sigma_p)$ (for all σ in $\mathcal{E}^*(M)$) and mimicking the proof of the Leibniz rule for functions of one real variable. □

Lemma. *Let M be a manifold, X in $\mathfrak{X}(M)$ and f in $\mathcal{E}^0(M) = \mathcal{E}(M)$. Then*

$$\mathcal{L}_X f = X(f), \quad \textit{i.e.} \quad \mathcal{L}_X f = D^X(f),$$

where D^X is the derivation of $\mathcal{E}(M)$ canonically associated to X.

If furthermore η is in $\mathcal{E}^k(M)$ with $k\geq 1$ we have $\mathcal{L}_X(f\cdot\eta) = (\mathcal{L}_X f)\cdot\eta + f\cdot(\mathcal{L}_X\eta)$.

Proof. Let p be in M, then

$$(\mathcal{L}_X f)(p) = \frac{d}{dt}\Big|_0((\varphi_t^X)^*f)_p = \frac{d}{dt}\Big|_0(f\circ\varphi_t^X)(p) = \frac{d}{dt}\Big|_0 f(\varphi_t^X(p)),$$

i.e. $(\mathcal{L}_X f)(p)$ is the derivative of f in the direction $\frac{d}{dt}|_0\varphi_t^X(p) = X(p)$. Since this is the very definition of $(D^X(f))(p)$ the first part is proven.

The second part follows as sketched in the proof of Assertion (iv) of the preceding proposition. □

Definition. Let M be a manifold and X and Y in $\mathfrak{X}(M)$. The "Lie derivative of Y with respect to X" is defined as follows

$$(\mathcal{L}_X Y)(p) := \frac{d}{dt}\Big|_0\{((\varphi_{-t}^X)_*)_{\varphi_t^X(p)}((\varphi_t^X(p)))\} \quad \text{for all } p \text{ in } M.$$

Remark. Since the flow φ^X is smooth $\mathcal{L}_X Y$ is a well–defined section of TM. Its smoothness will follow a fortiori from the next proposition.

Definition. Let M' and M'' be manifolds and $F : M' \to M''$ a diffeomorphism. For a vector field Z on M' we define its "push–forward" as follows

$$(F_*Z)_q := (T_{F^{-1}(q)}F)(Z_{F^{-1}(q)}) = (F_*)_{F^{-1}(q)}(Z_{F^{-1}(q)}) \quad \text{for all } q \text{ in } M''.$$

Remark. Considering a fixed point $q = F(p)$ the above formula is just the tangent of the map F. In the case here considered when F is a diffeomorphism, F_*Z is easily seen to be a smooth section of TM'', i.e. F_*Z is a vector field on M''.

Lemma. *Let M', M'', F and Z be as in the preceding definition, and let φ_t^Z be the flow of Z on M'. Then the flow of F_*Z on M'' is given by $\varphi_t^{F_*Z} = F \circ \varphi_t^Z \circ F^{-1}$.*

Furthermore, if f is in $\mathcal{E}(M'')$ then $((F_*Z)(f))(q) = (Z(f \circ F))(F^{-1}(q))$ for all q in M''. ($(F_*Z)(f)$ denotes of course again $D^{F_*Z}(f)$).

Proof. Obviously we have $F \circ \varphi_0^Z \circ F^{-1} = \mathrm{Id}_{M''}$. Furthermore

$$\left.\frac{d}{dt}\right|_0 (F \circ \varphi_t^Z \circ F^{-1})(q) = (T_{F^{-1}(q)}F)(Z_{F^{-1}(q)}) = (F_*Z)_q$$

so that the first assertion is proven.

Let now f be in $\mathcal{E}(M'')$ then

$$((F_*Z)(f))(q) = \left.\frac{d}{dt}\right|_0 (f \circ \varphi_t^{F_*Z})(q) = \left.\frac{d}{dt}\right|_0 (f \circ F \circ \varphi_t^Z \circ F^{-1})(q)$$

$$= \left.\frac{d}{dt}\right|_0 (f \circ F)(\varphi_t^Z(F^{-1}(q))) = Z_{F^{-1}(q)}(f \circ F) = (Z(f \circ F))(F^{-1}(q)).$$

□

Proposition. *Let M be a manifold and X and Y in $\mathfrak{X}(M)$. Then*

$$\mathcal{L}_X Y = [X, Y].$$

Proof. Since $\mathfrak{X}(M)$ is canonically isomorphic to the derivations of $\mathcal{E}(M)$ it is enough to show that $(\mathcal{L}_X Y)(f) = ([X, Y])(f)$ for all f in $\mathcal{E}(M)$.

Let f be in $\mathcal{E}(M)$ and p a point in M. There is a $\epsilon = \epsilon(p) > 0$ and an open neighborhood U of p in M such that $\varphi_t^X|_U : U \to W_t = \varphi_t^X(U)$ is a diffeomorphism for all t with $|t| < \epsilon$. (The sets W_t are of course open in M.)

We compute

$$(\mathcal{L}_X Y)_p(\underset{\sim p}{f}) = \left.\frac{d}{dt}\right|_0 \{((\varphi_{-t}^X)_*\Big)_{\varphi_t^X(p)}(Y(\varphi_t^X(p)))\}(\underset{\sim p}{f}) = \left.\frac{d}{dt}\right|_0 \{((\varphi_{-t}^X)_* Y)_p(\underset{\sim p}{f})\},$$

where we consider for every t the diffeomorphism $F = \varphi_{-t}^X : W_t = M' \to U = M''$ and we use the pushforward notation.

By the previous lemma we have

$$(\mathcal{L}_X Y)_p(\underset{\sim p}{f}) = \left.\frac{d}{dt}\right|_0 \{(Y(f \circ \varphi_{-t}^X))(\varphi_t^X(p))\}.$$

In order to calculate this derivative with respect to t we introduce the smooth maps

$$\Delta : I_\epsilon := (-\epsilon, \epsilon) \to I_\epsilon \times I_\epsilon, \Delta(t) = (\Delta_1(t), \Delta_2(t)) := (t, t) \quad \text{and}$$

$$\psi : I_\epsilon \times I_\epsilon \to \mathbb{R}, \psi(r, s) = (Y(f \circ \varphi_{-s}^X))(\varphi_r^X(p)).$$

(If necessary we choose a smaller $\epsilon > 0$ here.)

It follows that

$$(\mathcal{L}_X Y)_p = \frac{d}{dt}\Big|_0 (\psi \circ \Delta)(t) = \Big(\frac{\partial \psi}{\partial r}(0,0)\Big) \cdot \Big(\frac{d\Delta_1}{dt}(0)\Big) + \Big(\frac{\partial \psi}{\partial s}(0,0)\Big) \cdot \Big(\frac{d\Delta_2}{dt}(0)\Big)$$
$$= (X(Y(f))(p)) \cdot 1 - (Y(X(f))(p)) \cdot 1 = ([X,Y](f))(p).$$

□

Proposition. *Let M be a manifold, X a vector field on M and η a k–form on M (with $k \geq 1$). Then*

$$(\mathcal{L}_X \eta)(X_1, \ldots, X_k) = \mathcal{L}_X(\eta(X_1, \ldots, X_k)) - \sum_{j=1}^{k} \eta(X_1, \ldots, X_{j-1}, \mathcal{L}_X X_j, X_{j+1}, \ldots, X_k)$$

for all vector fields $X_1, \ldots, X_k$ on M .

Proof. For all (t,p) in Ω, the domain of the local flow of X, we have

$$((\varphi_t^X)^*(\eta(X_1, \ldots, X_2)))(p) = \{((\varphi_{-t}^X)^*((\varphi_t^X)^*\eta))(X_1, \ldots, X_k)\}(\varphi_t^X(p))$$
$$= ((\varphi_{-t}^X)^*((\varphi_t^X)^*\eta))_{\varphi_t^X(p)}\Big(X_1(\varphi_t^X(p)), \ldots, X_k(\varphi_t^X(p))\Big)$$
$$= ((\varphi_t^X)^*\eta)_p\Big(((\varphi_{-t}^X)_*)_{\varphi_t^X(p)}(X_1(\varphi_t^X(p))), \ldots, ((\varphi_{-t}^X)_*)_{\varphi_t^X(p)}(X_k(\varphi_t^X(p)))\Big).$$

Thus, using the Leibniz rule repeatedly in the variable t as in the preceding proof one has

$$(\mathcal{L}_X(\eta(X_1, \ldots, X_k)))(p) = \frac{d}{dt}\Big|_0 ((\varphi_t^X)^*(\eta(X_1, \ldots, X_k)))(p)$$
$$= ((\mathcal{L}_X \eta)(X_1, \ldots, X_k))(p) + \sum_{j=1}^{k} \eta(X_1, \ldots, X_{j-1}, \mathcal{L}_X X_j, X_{j+1}, \ldots, X_k)(p).$$

□

Bibliographical remarks. *Books on manifolds as quoted at the end of 2.1 plus [GHL].*

2.7 The exterior derivative of differential forms and de Rham cohomology

Example. Let Ω be an open subset of $\mathbb{R}^3$. A vector field, given as $K = \sum_{j=1}^{3} K_j \frac{\partial}{\partial x_j}$ with K_j in $\mathcal{E}(\Omega)$, is – especially in physics – often described by the smooth vector–valued function $\vec{K} = {}^t(K_1, K_2, K_3) : \Omega \to \mathbb{R}^3$. The canonical Riemannian structure and orientation allow the definition of the usual operations of vector calculus (for ψ in $\mathcal{E}(\Omega)$ the expression $\partial_j\psi$ denotes here the partial derivative $\frac{\partial\psi}{\partial x_j}$ with j in $\{1,2,3\}$):

* the "gradient" of a smooth real–valued function f on Ω is defined as

$$\operatorname{grad} f = \nabla f = \sum_{j=1}^{3} \partial_j f \frac{\partial}{\partial x_j}$$

or by $\vec{\nabla} f = {}^t(\partial_1 f, \partial_2 f, \partial_3 f)$,

* the "curl" of a vector field K is given as

$$\operatorname{curl} K = \nabla \times K = (\partial_2 K_3 - \partial_3 K_2)\frac{\partial}{\partial x_1} + (\partial_3 K_1 - \partial_1 K_3)\frac{\partial}{\partial x_2} + (\partial_1 K_2 - \partial_2 K_1)\frac{\partial}{\partial x_3}$$

or as $\vec{\nabla} \times \vec{K}$, where $\vec{\nabla} = {}^t(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3})$ and "$\times$" is the vector product,

* the "divergence" of a vector field L is

$$\operatorname{div} L = \sum_{j=1}^{3} \partial_j L_j$$

or as $\vec{\nabla} \cdot \vec{L}$, where "$\cdot$" is the scalar product.

It is easily checked that $\operatorname{div}(\operatorname{curl} K) = 0$ and $\operatorname{curl}(\operatorname{grad} f) = 0$. Thus in order to decide e.g. if a given force field $\vec{K}$ is conservative, i.e., a gradient field, we immediately find the necessary condition $\operatorname{curl}(K) = 0$.

In order to translate the "exact sequence" (i.e., $\operatorname{div} \circ \operatorname{curl} = 0$ and $\operatorname{curl} \circ \operatorname{grad} = 0$)

$$\mathcal{E}^0(\Omega) \xrightarrow{\text{grad}} \mathfrak{X}(\Omega) \xrightarrow{\text{curl}} \mathfrak{X}(\Omega) \xrightarrow{\text{div}} \mathcal{E}^0(\Omega)$$

into the language of differential forms we define the following $\mathcal{E}(\Omega)$–module isomorphisms

$$\begin{aligned}
\tau_1 \ &: \ \mathfrak{X}(\Omega) \to \mathcal{E}^1(\Omega), \ \tau_1(\sum_{j=1}^{3} K_j \frac{\partial}{\partial x_j}) = \sum_{j=1}^{3} K_j \, dx_j; \\
\tau_2 \ &: \ \mathfrak{X}(\Omega) \to \mathcal{E}^2(\Omega), \ \tau_2(\sum_{j=1}^{3} L_j \frac{\partial}{\partial x_j}) \\
&\qquad = L_1 dx_2 \wedge dx_3 - L_2 dx_1 \wedge dx_3 + L_3 dx_1 \wedge dx_2, \\
\tau_3 \ &: \ \mathcal{E}^0(\Omega) \to \mathcal{E}^3(\Omega), \ f \mapsto f\, dx_1 \wedge dx_2 \wedge dx_3.
\end{aligned}$$

Setting $\tilde{d}_0 := \tau_1 \circ \operatorname{grad}$, $\tilde{d}_1 := \tau_2 \circ \operatorname{curl} \circ (\tau_1)^{-1}, \tilde{d}_3 := \tau_3 \circ \operatorname{div} \circ (\tau_2)^{-1}$ we get an exact sequence ($\tilde{d}_{j+1} \circ \tilde{d}_j = 0$) as follows:

$$\mathcal{E}^0(\Omega) \xrightarrow{\tilde{d}_0} \mathcal{E}^1(\Omega) \xrightarrow{\tilde{d}_1} \mathcal{E}^2(\Omega) \xrightarrow{\tilde{d}_2} \mathcal{E}^3(\Omega).$$

Over first objective is to generalize this sequence to an arbitrary manifold.

Definition. Let M be a manifold and η a differential k–form on M ($k \geq 0$). The "exterior derivative $d(\eta)$ of η" is defined by the following formula:

$$(d(\eta))(X_1, \ldots, X_{k+1}) := \sum_{j=1}^{k+1} (-1)^{j+1} X_j(\eta(X_1, \ldots, \hat{X}_j, \ldots, X_{k+1}))$$

$$+ \sum_{i<j} (-1)^{i+j} \eta([X_i, X_j], X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{k+1})$$

for all $X_1, \ldots, X_{k+1}$ in $\mathfrak{X}(M)$. (*A* hat "$\wedge$" on a vector field means that the corresponding vector field is omitted.)

Remarks. (1) The expression $d(\eta)$ defines a smooth $(k+1)$–form on M. (Proof as an exercise.)

(2) For f in $\mathcal{E}(M) = \mathcal{E}^0(M)$ and X in $\mathfrak{X}(M)$ we have $(d(f))(X) = X(f) = D^X(f)$, and for α in $\mathcal{E}^1(M)$ and X_1, X_2 in $\mathfrak{X}(M)$ we have

$$(d(\alpha))(X_1, X_2) = X_1(\alpha(X_2)) - X_2(\alpha(X_1)) - \alpha([X_1, X_2]).$$

Let us recall that a k–form η on an open set $V \subset \mathbb{R}^m$ can be described as a finite sum $\sum\limits_{|J|=k} f_J dx_J$, where $J = (j_1, \ldots, j_k)$ is a multi–index of "length k" ($|J| = k$), $j_1 < \cdots < j_k$, and $dx_J = dx_{j_1} \wedge \ldots \wedge dx_{j_k}$ and f_J in $\mathcal{E}^0(V)$.

Lemma. *Let f, f_J be in $\mathcal{E}^0(V)$ for V an open set in $\mathbb{R}^m$ and $\eta = \sum\limits_{|J|=k} f_J dx_J$ a differential k–form with $k \geq 1$. Then*

(i) $d(f) = \sum\limits_{j=1}^{m} \frac{\partial f}{\partial x_j} dx_j$, and

(ii) $d(\eta) = \sum\limits_{|J|=k} (d(f_J)) \wedge dx_J$.

Proof. *Ad (i).* Since *df* is a 1–form there are function g_j on V such that $d(f) = \sum_{j=1}^m g_j dx_j$. We calculate

$$g_l = \Big(\sum_{j=1}^{m} g_j dx_j\Big)(\frac{\partial}{\partial x_l}) = (d(f))(\frac{\partial}{\partial x_l}) = (\frac{\partial}{\partial x_l})(f) = \frac{\partial f}{\partial x_l}.$$

Ad(ii). Since the operator $d : \mathcal{E}^k(M) \to \mathcal{E}^{k+1}(M)$ is obviously $\mathbb{R}$–linear we may assume that $\eta = f\, dx_1 \wedge \ldots \wedge dx_k$. Since $d(\eta)$ is a $(k+1)$–form we know that

$$d(\eta) = \sum_{l=k+1}^{m} g_l dx_1 \wedge \ldots \wedge dx_k \wedge dx_l + \sum_{|J'|=k+1} g_{J'}, dx_{J'},$$

where $J' = (j'_1, \ldots, j'_{k+1})$ with $j'_1 < \cdots < j'_{k+1}$ and $\{1, \ldots, k\} \not\subset \{j'_1, \ldots, j'_{k+1}\}$. We have

$$g_{J'} = (d(\eta))(\frac{\partial}{\partial x_{j'_1}}, \ldots, \frac{\partial}{\partial x_{j'_{k+1}}}) = \sum_{r=1}^{k+1} (-1)^{r+1} \frac{\partial}{\partial x_{j'_r}} \Big(\eta(\frac{\partial}{\partial x_{j'_1}}, \ldots, \widehat{\frac{\partial}{\partial x_{j'_r}}}, \ldots, \frac{\partial}{\partial_{j'_{k+1}}})\Big)$$

$$+\sum_{r<s}(-1)^{r+s}\eta\Big(\left[\frac{\partial}{\partial x_{j'_r}},\frac{\partial}{\partial x_{j'_s}}\right],\frac{\partial}{\partial x_{j'_1}},\ldots,\widehat{\frac{\partial}{\partial x_{j'_r}}},\ldots,\widehat{\frac{\partial}{\partial x_{j'_s}}},\ldots,\frac{\partial}{\partial x_{j'_{k+1}}}\Big)=0$$

since by $[\frac{\partial}{\partial x_i},\frac{\partial}{\partial x_j}]=0$ the second sum vanishes and since the first sum is zero by $J'\backslash\{j'_r\}\neq\{1,\ldots,k\}$ for all r.

Furthermore

$$g_l=(d(\eta))(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_k},\frac{\partial}{\partial x_l})=\sum_{i=1}^{k}(-1)^{i+1}\frac{\partial}{\partial x_i}\Big(\eta(\frac{\partial}{\partial x_1},\ldots,\widehat{\frac{\partial}{\partial x_i}},\ldots\frac{\partial}{\partial x_k},\frac{\partial}{\partial x_l})\Big)$$

$$+(-1)^{k+1+1}\frac{\partial}{\partial x_l}\Big(\eta(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_k})\Big)+\sum(\pm1)\eta\Big(\left[\frac{\partial}{\partial x_r},\frac{\partial}{\partial x_s}\right],\frac{\partial}{\partial x_t},\ldots\Big).$$

The third sum vanishes termwise as above since the "coordinate vector fields commute", i.e. $\left[\frac{\partial}{\partial x_r},\frac{\partial}{\partial x_s}\right]=0$ for all r and s. The first sum also vanishes termwise since $(dx_1\wedge\ldots\wedge dx_k)(\frac{\partial}{\partial x_1},\ldots,\widehat{\frac{\partial}{\partial x_i}},\ldots,\frac{\partial}{\partial x_k},\frac{\partial}{\partial x_l})$ is zero for all $l>k$. Thus we find

$$g_l=(-1)^{k+2}\frac{\partial}{\partial x_l}(f)$$

and hence

$$d(\eta)=\sum_{l=k+1}^{m}(-1)^{k+2}\frac{\partial f}{\partial x_l}dx_1\wedge\ldots\wedge dx_k\wedge dx_l=\sum_{l=k+1}^{m}\frac{\partial f}{\partial x_l}dx_l\wedge dx_1\wedge\ldots\wedge dx_k$$

$$=\sum_{l=1}^{m}\frac{\partial f}{\partial x_l}dx_l\wedge dx_1\wedge\ldots\wedge dx_k=(d(f))\wedge dx_1\wedge\ldots\wedge dx_k.$$

□

Lemma. *Let V be open in $\mathbb{R}^m$ and j in $\{1,\ldots,m\}$. Then*

$$d(x_j)=dx_j.$$

Proof. The right hand side is the 1–form uniquely determined by $(dx_j)(\frac{\partial}{\partial x_l})=\delta_{j,l}$ for $l=1,\ldots,m$. Let us compare this to the exterior derivative of the j–th coordinate function

$$(d(x_j))(\frac{\partial}{\partial x_l})=\frac{\partial}{\partial x_l}(x_j)=\delta_{j,l}.$$

□

Remark and definition. Since the preceding lemma shows that there is no difference between $d(x_j)$ and dx_j from now we will simplify the notation and will not distinguish between them, i.e.

$$d\eta:=d(\eta)\quad\text{for all }\eta\text{ in }\quad\mathcal{E}^*(M).$$

Proposition. *Let M be a manifold of dimension m, η a k–form and μ a l–form on M. Then*

(i) $d(\lambda \cdot \eta) = \lambda \cdot d\eta$ *for* λ *in* $\mathbb{R}$,

(ii) $d(\eta + \mu) = d\eta + d\mu$ *if* $k = l$,

(iii) $d(\eta \wedge \mu) = (d\eta) \wedge \mu + (-1)^k \eta \wedge d\mu$,

(iv) $d\eta = 0$ *if* $k \geq m$,

(v) $d(d\eta) = 0$.

Proof. The first two assertions follow directly from the definition and the fourth one from the fact that there are no non–zero n–forms for $n > m$ on a m–dimensional manifold.

In order to show the remaining assertions we observe first that d is a "local operator", i.e. to calculate $(d\eta)_p$ it is enough to consider η on an open neighborhood of p. We can therefore assume for the rest of the proof that $M = V$ is an open subset of $\mathbb{R}^m$.

Furthermore we may assume that $\eta = f\,dx_{i_1} \wedge \ldots \wedge dx_{i_k}$ and $\mu = g\,dx_{j_1} \wedge \ldots \wedge dx_{j_l}$.

It follows

$$d(\eta \wedge \mu) = d(fg\,dx_{i_1} \wedge \ldots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_l})$$

$$= d(fg) \wedge dx_{i_1} \wedge \ldots \wedge dx_{j_l} = (g(df) + f(dg)) \wedge dx_{i_1} \wedge \ldots \wedge dx_{j_l} = (df \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_k})$$

$$\wedge (g\,dx_{j_1} \wedge \ldots \wedge dx_{j_l}) + (-1)^k (f\,dx_{i_1} \wedge \ldots \wedge dx_{i_k}) \wedge (dg \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_l}) = (d\eta) \wedge \mu + (-1)^k \eta \wedge (d\mu).$$

Ad(iv). Let f be in $\mathcal{E}^0(M)$ and X_1, X_2 in $\mathfrak{X}(M)$. Then

$$(d(df))(X_1, X_2) = X_1((df)(X_2)) - X_2((df)(X_1)) - (df)([X_1, X_2])$$

$$= X_1(X_2(f)) - X_2(X_1(f)) - [X_1, X_2](f) = 0.$$

By localizing near a point p in M we can again assume that $M = V$, open in $\mathbb{R}^m$, and that $\eta = f\,dx_{i_1} \wedge \ldots \wedge\, dx_{i_k}$. It follows that

$$d(d\eta)) = d((df) \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_k})$$

$$= (d(df)) \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_k} + (df) \wedge d(dx_{i_1} \wedge \ldots \wedge dx_{i_k}) = 0$$

since $d^2(f) = 0$ and $d(dx_i) = 0$. □

Example (and exercise). Check that in the example at the beginning of this section the operators $\tilde{d}_j$ are the operator d in the different degrees (of differential forms.) Translate the assertions of the preceding proposition to formulas in vector analysis.

Remark. Since $\mathcal{E}^*(M) = \oplus_{k \geq 0} \mathcal{E}^k(M)$ the operators $d : \mathcal{E}^k(M) \to \mathcal{E}^{k+1}(M)$ can be put together to build an operator $d : \mathcal{E}^*(M) \to \mathcal{E}^*(M)$. The assertions (i) – (iii) can then be rephrased by saying that d is an odd super–derivation of the real super–algebra $\mathcal{E}^*(M)$.

Proposition. *Let M and N be manifolds, and d_M and d_N the respective exterior derivatives. Let furthermore $F : M \to N$ be a smooth map then*

$$d_M \circ F^* = F^* \circ d_N.$$

Proof. Going to local coordinates we may assume that $M = V$ is open in $\mathbb{R}^m$, $N = W$ is open in $\mathbb{R}^n$, and that the map $F : M \to N$ is given by n scalar functions $F = {}^t(F_1, \ldots, F_n)$.

Given a k–form η on W we can assume that $\eta = g\, dy_1 \wedge \ldots \wedge dy_k$, where g is in $\mathcal{E}^0(W)$ and $y_1, \ldots, y_k$ are the first k coordinate functions on $W \subset \mathbb{R}^n$.

We calculate (setting intermediately $d_M = d_N = d$ for convenience):

$$(d_M \circ F^*)(\eta) = d_M(F^*(g\, dy_1 \wedge \ldots \wedge dy_k)) = d_M((g \circ F) dF_1 \wedge \ldots \wedge dF_k)$$

$$= d(g \circ F) \wedge dF_1 \wedge \ldots \wedge dF_k = \sum_{j=1}^{n} \sum_{i=1}^{m} (\frac{\partial g}{\partial y_j} \circ F) \frac{\partial F_j}{\partial x_i} dx_i \wedge dF_1 \wedge \ldots \wedge dF_k$$

$$= \Big(\sum_{j=1}^{n} F^*(\frac{\partial g}{\partial y_j}) dF_j \Big) \wedge dF_1 \wedge \ldots \wedge dF_k = F^*(dg) \wedge dF_1 \wedge \ldots \wedge dF_k = (F^* \circ d_N)(\eta).$$

□

Corollary. *Let M be a manifold and X a vector field on M. Then the following identity of operators holds on $\mathcal{E}^*(M)$.*

$$\mathcal{L}_X \circ d = d \circ \mathcal{L}_X.$$

Proof. Since $(\mathcal{L}_X \eta) = \frac{d}{dt}\Big|_0 ((\varphi_t^X)^* \eta)$ the assertion follows easily from the preceding proposition. □

Proposition ("Cartan's magic formula"). *Let M be a manifold and X a vector field on M. Then the following identity of operators holds on $\mathcal{E}^*(M)$:*

$$\mathcal{L}_X = d \circ i_X + i_X \circ d.$$

Proof. The formula holds trivially for 0–forms, i.e. functions on M. Let now $k \geq 1$ and η a k–form on M. Let furthermore $X_1, \ldots, X_k$ be in $\mathfrak{X}(M)$ and set $X_0 = X$. Then

$$((d \circ i_X + i_X \circ d)(\eta))(X_1, \ldots X_k) = (d(\eta(X, \ldots)))(X_1, \ldots, X_k) + ((d\eta)(X_0, \ldots))(X_1, \ldots, X_k)$$

$$= \sum_{i=1}^{k} (-1)^{i+1} X_i(\eta(X_0, X_1, \ldots, \widehat{X}_i, \ldots, X_k))$$

$$+ \sum_{1 \leq i < j} (-1)^{i+j} \eta(X_0, [X_i, X_j], X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_k) + (-1)^{0+1+1} X_0(\eta(X_1, \ldots, X_k))$$

$$+ \sum_{i=1}^{k} (-1)^{i+2} X_i(\eta(X_0, \ldots, \widehat{X}_i, \ldots, X_k)) + \sum_{0 < j} (-1)^j \eta([X_0, X_j], X_1, \ldots, \widehat{X}_j, \ldots, X_k)$$

$$+ \sum_{0 < i < j} (-1)^{i+j} \eta([X_i, X_j], X_0, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_k)$$

$$= X_0(\eta(X_1,\dots,X_k)) - \sum_{0<j} \eta(X_1,\dots,X_{j-1},[X_0,X_j],X_{j+1},\dots,X_k)$$
$$= (\mathcal{L}_X\eta)(X_1,\dots,X_k).$$

□

Remark. Beside this "algebraic" proof of Cartan's formula there is also an "analytic" proof using the flow of the vector field X. We will not present it here but we will use the latter approach to get a stronger result for "time–dependent vector fields" as a preparation for Moser's method to prove Darboux's theorem.

Corollary. *Let M be a manifold, X, Y in $\mathfrak{X}(M)$ and f in $\mathcal{E}^0(M)$. Then for η in $\mathcal{E}^k(M)$ one has*

(i) $\mathcal{L}_{fX}\eta = f\cdot\mathcal{L}_X\eta + (df)\wedge(i(X)\eta)$.

Furthermore the following identies hold on $\mathcal{E}^*(M)$

(ii) $[\mathcal{L}_X, i_Y] = i_{[X,Y]}$,

(iii) $[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_{[X,Y]}$.

Proof. *Ad(i).*

$$\mathcal{L}_{fX}\eta = d(\eta(fX,\dots)) + (d\eta)(f\cdot X,\dots) = d(f\cdot\eta(X,\dots)) + f\cdot(d\eta)(X,\dots)$$
$$= df\wedge(\eta(X,\dots)) + f\cdot d(\eta(X,\dots)) + f\cdot(d\eta)(X,\dots)$$
$$= (df)\wedge(i(X)\eta) + f\cdot(d\circ i_X + i_X\circ d)(\eta) = f\cdot\mathcal{L}_X\eta + (df)\wedge(i(X)\eta).$$

The second assertion follows from the explicit formula for $\mathcal{L}_Z\mu$, $Z\in\mathfrak{X}(M)$, μ a differential form shown at the end of Section 2.6.

The third formula is then derived from the second and Cartan's magic formula. □

Remark. If M is a manifold of dimension m then we have the sequence

$$\mathcal{E}^0(M)\xrightarrow{d}\mathcal{E}^1(M)\xrightarrow{d}\cdots\xrightarrow{d}\mathcal{E}^k(M)\xrightarrow{d}\mathcal{E}^{k+1}(M)\xrightarrow{d}\cdots\xrightarrow{d}\mathcal{E}^m(M)\xrightarrow{d}0,$$

which is "exact" in the sense that $d\circ d = 0$.

Definition. Let M be a manifold, then for $k\geq 1$ the "k–th de Rham–cohomology group of M" is defined as follows

$$H^k_{dR}(M) := H^k_{dR}(M,\mathbb{R}) := \frac{\ker(d:\mathcal{E}^k(M)\to\mathcal{E}^{k+1}(M))}{\operatorname{im}(d:\mathcal{E}^{k-1}(M)\to\mathcal{E}^k(M))},$$

where ker and im denote the kernel and the image of the corresponding $\mathbb{R}$–linear maps.

For $k = 0$ one defines

$$H^0_{dR}(M) := H^0_{dR}(M,\mathbb{R}) := \ker(d:\mathcal{E}^0(M)\to\mathcal{E}^1(M)).$$

Remarks. (1) A differential form η such that $d\eta = 0$ is called "closed", and if there is a form μ fulfilling $d\mu = \eta$ we call η "exact". The Rham cohomology groups measure therefore "how many closed forms on M are not exact."

(2) Though traditionally called "cohomology groups" the spaces $H^k_{dR}(M)$ are in fact $\mathbb{R}$–vector spaces.

(3) If the dimension of M is less or equal than m, then $H^k_{dR}(M) = \{0\}$ for all $k \geq m+1$.

(4) A function f in $\mathcal{E}^0(M)$ is in the kernel of d if only if f is locally constant. (Proof as an exercise.) Thus for a connected manifold M we have that $H^0_{dR}(M) \cong \mathbb{R}$.

(5) Tensorizing the bundles $\Lambda^k T^*M$ with the trivial complex bundle $M \times \mathbb{C} \to \mathbb{C}$ we get smooth complex vector bundles $\Lambda^k T^*M \bigotimes_{\mathbb{R}} \mathbb{C} \longrightarrow M$. Its sections $\Gamma_{C^\infty}(M, \Lambda^k T^*M \bigotimes_{\mathbb{R}} \mathbb{C})$ are denoted by $\mathcal{E}^k_{\mathbb{C}}(M)$ and called "complex(–valued) differential k–forms on M". The exterior derivative d can be extended to the spaces $\mathcal{E}^k_{\mathbb{C}}(M)$ by complex–linearity and we can, in complete analogy to the preceding definition, define $H^k_{dR}(M, \mathbb{C})$. It is not difficult to see that $H^k_{dR}(M, \mathbb{C}) \cong H^k_{dR}(M, \mathbb{R}) \bigotimes_{\mathbb{R}} \mathbb{C}$.

Remark. Though easy to define the de Rham cohomology groups of a given manifold M might be difficult to calculate.

Example (and exercise). Let again Ω be open in $\mathbb{R}^3$ and $\vec{K} : \Omega \to \mathbb{R}^3$ a smooth force field on Ω. Recall that $\vec{K}$ is called conservative if and only if there exists a smooth function $V : \Omega \to \mathbb{R}$, a potential, such that $\vec{K} = \vec{\nabla} V$. Since curl $\circ$ grad $= 0$ (i.e. $d^2 = 0$), we arrive at the necessary condition $\vec{\nabla} \times \vec{K} = 0$, which is equivalent to $d(\sum_{j=1}^3 K_j dx_j) = 0$. Thus for Ω open in $\mathbb{R}^3$ we have

$$H^1_{dR}(\Omega) \cong \frac{\{\text{curl-free force fields on } \Omega\}}{\{\text{conservative force fields on } \Omega\}}.$$

Let now $\Omega := \mathbb{R}^3 \backslash \{{}^t(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1 = 0 \text{ and } x_2 = 0\}$ and $\vec{K} = {}^t\left(\frac{-x_2}{x_1^2+x_2^2}, \frac{x_1}{x_1^2+x_2^2}, 0\right)$.

A direct calculation shows that $\vec{\nabla} \times \vec{K} = 0$, i.e. $\vec{K}$ is curl-free.

Assuming now that $\vec{K} = \vec{\nabla} V$ for a function $V : \Omega \to \mathbb{R}$, then the work along a path should depend only on the endpoints, i.e. in physicists' language

$$\int_C \vec{K} \cdot \vec{ds} = V(b) - V(a)$$

if C is a path from a to b.

Defining now paths C_ϵ for $\epsilon = \pm 1$ by maps $\gamma_\epsilon : [0,1] \to \Omega$ as follows

$$\gamma_\epsilon(t) := \begin{pmatrix} \cos(\pi t) \\ \epsilon \cdot \sin(\pi t) \\ 0 \end{pmatrix}$$

we find $\gamma_\epsilon(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \gamma_\epsilon(1) = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$ and $\int_{C_\epsilon} \vec{K} \cdot \vec{ds} = \pi\epsilon$, i.e. $\vec{K}$ cannot be a conservative field.

Otherwise stated, the differential 1–form $\alpha_K = \sum_{j=1}^3 K_j dx_j$ defines a non–zero class in $H^1_{dR}(\Omega)$. In fact, one has $H^1_{dR}(\Omega) = (([\alpha_K]))_{\mathbb{R}}$.

In order to prove the aforementioned fact as well as the "Poincaré lemma" below we will first investigate the relation between homotopies and cohomology in general.

Definitions. Let M and N be manifolds and A a closed submanifold of M.

(1) For a smooth map $F : [0,1] \times M \to N$ we set $F_t(p) := F(t,p)$. We call F a "(smooth) homotopy between the maps F_0 and F_1 (from M to N)."

(2) Let $f : M \to N$ and $g : N \to M$ be smooth maps such that $g \circ f$ is homotopic to Id_M and $f \circ g$ homotopic to Id_N. We then say that "M and N have the same homotopy type (in the C^∞–sense)".

(3) If M has the same homotopy type as a point we call M "contractible".

(4) If $i : A \to M$ is the inclusion map and $r : M \to A$ is a smooth map that restricts to the identity on A, i.e. $r \circ i = \mathrm{Id}_A$, we call r a "retraction of M to A". If furthermore $i \circ r : M \to M$ is homotopic to the identity of M we call r a "deformation retraction of M onto A."

Remark. A map $F : [0,1] \times M \to N$ is smooth if F_t is smooth for all t and if $t \mapsto F(t,p)$ is smooth for all p in M, where smoothness in the boundary points is defined by taking appropriate (one–sided) differential quotients.

Proposition. *Let $F : [0,1] \times M \to N$ be a smooth map. Then there exists a $\mathbb{R}$–linear operator $H : \mathcal{E}^*(N) \to \mathcal{E}^*(M)$, lowering the degree of differential forms by one, that satisfies*

$$d \circ H + H \circ d = F_1^* - F_0^*.$$

Remark. Such an operator H is called a "homotopy operator".

Proof of the proposition.

Let μ be in $\mathcal{E}^l(N)$, p in M and $v_1, \dots, v_{l-1}$ in T_pM. We set for t in $[0,1]$: $i_t : M \to [0,1] \times M, i_t(p) = (t,p)$ and define

$$(H\mu)_p(v_1, \dots, v_{l-1}) := \int_0^1 \Big[(F^*\mu)_{(t,p)}\big(\frac{\partial}{\partial t}\Big|_{(t,p)}, (i_t)_{*p}v_1, \dots, (i_t)_{*p}v_{l-1}\big) \Big]\, dt.$$

(Since $[0,1] \times M$ is a product the injections i_t are obvious and we will omit the maps $(i_t)_*$ in the rest of the proof.)

In order to show the asserted formula it is enough to consider an open neighborhood of a given point p in M, i.e. we can assume that $M = V$ is open in $\mathbb{R}^m$. Thus $H\mu$ is determined by its values on the "coordinate fields":

$$(H\mu)_x\big(\frac{\partial}{\partial x_{i_1}}\Big|_x, \dots, \frac{\partial}{\partial x_{i_{l-1}}}\Big|_x\big) = \int_0^1 \Big[(F^*\mu)_{(t,x)}\big(\frac{\partial}{\partial t}\Big|_{(t,x)} \frac{\partial}{\partial x_{i_1}}\Big|_{(t,x)}, \dots, \frac{\partial}{\partial x_{i_{l-1}}}\Big|_{(t,x)}\big) \Big]\, dt,$$

where x is in V and $i_1 < \dots < i_{l-1}$. This local description of $H\mu$ immediately shows that $H\mu$ is smooth in x as a "parameter–depending" integral.

Let us now calculate the terms in the left hand side of the assertion, applied to a fixed k–form η on N (and for $i_1 < \cdots < i_k$):

$$(d(H\eta))_x(\frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}}) = \sum_{j=1}^{k}(-1)^{j+1}(\frac{\partial}{\partial x_{i_j}}\Big|_x \Big((H\eta)(\frac{\partial}{\partial x_{i_1}}, \ldots, \widehat{\frac{\partial}{\partial x_{i_j}}}, \ldots, \frac{\partial}{\partial x_{i_k}})\Big) +$$

$$\sum_{r<s}(-1)^{r+s}(H\eta)_x\Big(\Big[\frac{\partial}{\partial x_{i_r}}, \frac{\partial}{\partial x_{i_s}}\Big], \ldots\Big)$$

$$= \sum_{j=1}^{k}(-1)^{j+1}\int_0^1 \left[\frac{\partial}{\partial x_{i_j}}\Big|_x ((F^*\eta)(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_{i_1}}, \ldots, \widehat{\frac{\partial}{\partial x_{i_j}}}, \ldots, \frac{\partial}{\partial x_{i_k}}))\right] dt,$$

since the coordinate fields commute. On the other hand

$$(H(d\eta))_x(\frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}}) = \int_0^1 \Big[(F^*d\eta)_x(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}})\Big] dt$$

$$= \int_0^1 \Big[(d(F^*\eta))_x(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}})\Big] dt = \int_0^1 \Big[\frac{\partial}{\partial t}\Big\{(F^*\eta)_x(\frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}})\Big\}\Big] dt$$

$$+ \sum_{j=1}^{k}(-1)^{j+2}\int_0^1 \Big[\frac{\partial}{\partial x_{i_j}}\Big|_x \Big\{(F^*\eta)(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_{i_1}}, \ldots, \widehat{\frac{\partial}{\partial x_{i_j}}}, \ldots, \widehat{\frac{\partial}{\partial x_{i_k}}})\Big\}\Big] dt,$$

since all commutator terms involue either $[\frac{\partial}{\partial t}, \frac{\partial}{\partial x_r}]$ or $[\frac{\partial}{\partial x_r}, \frac{\partial}{\partial x_s}]$ and thus vanish.

We therefore arrive at

$$(d(H\eta) + H(d\eta))_x(\frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}}) = \int_0^1 \Big[\frac{\partial}{\partial t}\Big\{(F^*\eta)_x(\frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}})\Big\}\Big] dt$$

$$= (F_1^*\eta - F_0^*\eta)_x(\frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}}).$$

□

Definition. Let M and N be manifolds and $f : M \to N$ be a smooth map. The "induced map on de Rham cohomology" is defined as follows

$$f^*([\eta]) := [f^*\eta] \quad \text{for all classes} \quad [\eta] \text{ in } H^*_{dR}(N).$$

Remark. Given a class c in $H^*_{dR}(N)$, one easily verifies that the class f^*c in $H^*_{dR}(M)$ is independent of its representative η in $\mathcal{E}^*(N)$.

Lemma. *Let M and N be manifolds and $f : M \to N$ a smooth map. Then*

(i) $f^* : H^*_{dR}(N) \to H^*_{dR}(M)$ is $\mathbb{R}$–linear,

(ii) $H^*_{dR}(M)$ is a super–commutative, associative real super–algebra with multiplication given by

$$[\eta] \wedge [\mu] := [\eta \wedge \mu],$$

(iii) f^* *is a algebra–homomorphism, i.e.* $[f^*\eta] \wedge [f^*\mu] = f^*[\eta \wedge \mu]$.

(iv) If L *is a further manifold and* $g : L \to M$ *is smooth then* $(f \circ g)^* = g^* \circ f^*$ *on de Rham cohomology.*

(v) If $M = N$ *and* $f = Id_M$ *then* $f^* = (Id_M)^* = Id_{H^*_{dR}(M)}$.

Proof. Exercise. □

We collect now several important applications of the last proposition (and the last lemma.)

Corollary 1. *Let* M *and* N *be manifolds and* $f, g : M \to N$ *be two smooth maps that are homotopic. Then*

$$f^* = g^* : H^*_{dR}(N) \to H^*_{dR}(M).$$

Proof. Let $F : [0,1] \times M \to N$ be a smooth map such that $F_0 = f$ and $F_1 = g$. For a class c in $H^k_{dR}(N)$ represented by a closed form η in $\mathcal{E}^k(N)$ we find

$$g^*c = g^*[\eta] = [g^*\eta] = [F_1^*\eta] = [F_0^*\eta + dH\eta + Hd\eta]$$

$$= [F_0^*\eta] + [dH\eta] = [F_0^*\eta] = [f^*\eta] = f^*[\eta] = f^*c.$$

□

Corollary 2. *Let* M *and* N *be manifolds having the same homotopy type. Then* $H^*_{dR}(M)$ *and* $H^*_{dR}(N)$ *are isomorphic as* $\mathbb{R}$*–algebras.*

Proof. Let $f : M \to N$ and $g : N \to M$ smooth maps such that $g \circ f$ is homotopic to Id_M and $f \circ g$ homotopic to Id_N. Then, by Corollary 1

$$f^* \circ g^* = (g \circ f)^* = (\mathrm{Id}_M)^* = \mathrm{Id}_{H^*_{dR}(M)} \quad \text{and}$$

$$g^* \circ f^* = (f \circ g)^* = (\mathrm{Id}_N)^* = \mathrm{Id}_{H^*_{dR}(N)}$$

and thus f^* and g^* are mutually inverse isomorphism between $H^*_{dR}(M)$ and $H^*_{dR}(N)$. □

Corollary 3. *Let* M *be a contractible manifold. Then* $H^0_{dR}(M) \cong \mathbb{R}$ *and* $H^k_{dR}(M) = \{0\}$ *for* $k > 0$.

Proof. Since the cohomology of the zero–dimensional connected manifold consisting of one point is isomorphic to $\mathbb{R}$ in degree zero and trivial in all other degrees the assertion follows from Corollary 2. □

Corollary 4. *Let* M *be a ball* $\mathbb{B}_R(0)$ *with radius* $0 < R \leq \infty$ *in* $\mathbb{R}^m$*, then* $H^0_{dR}(M) \cong \mathbb{R}$ *and* $H^k_{dR}(M) = \{0\}$ *for* $k > 0$.

Proof. The map

$$F : [0,1] \times \mathbb{B}_R(0) \to \mathbb{B}_R(0), F(t,x) = (1-t) \cdot x$$

is a smooth homotopy from M to the origin 0 in $\mathbb{B}_R(0)$ such that $F_t(0) = 0$ for all t. It follows easily that $\mathbb{B}_R(0)$ is contractible and thus by Corollary 3 the de Rham cohomology of $M = \mathbb{B}_R(0)$ is as asserted. □

Corollary 5. *Let M be a manifold and A a closed submanifold such there exists a deformation retraction $r : M \to A$ from M onto A. Then $r^* : H^*_{dR}(A) \to H^*_{dR}(M)$ is a $\mathbb{R}$-algebra isomorphism.*

Proof. Follows directly from Corollary 2. □

Corollary 6. *Let M and N be manifolds and $\varphi : M \to N$ a diffeomorphism. Then $\varphi^* : H^*_{dR}(N) \to H^*_{dR}(M)$ is an isomorphism.*

Proof. Let $\psi = \varphi^{-1} : N \to M$, then the assertions (iv) and (v) of the last lemma imply immediately that φ^* and ψ^* are mutually inverse isomorphisms in de Rham cohomology.

Examples (details as an exercise).

(1) Let M be $\mathbb{R}^m$. Then $H^0_{dR}(M) \cong \mathbb{R}$ and $H^k_{dR}(M) = \{0\}$ for $k > 0$.

(2) Let $E \xrightarrow{p} M$ be a vector bundle. Then $H^*_{dR}(M) \xrightarrow{p^*} H^*_{dR}(E)$ is an isomorphism.

(3) Let Ω be $\mathbb{R}^3 \backslash \{x_1 = 0, x_2 = 0\}$. Then $A = \{x_3 = 0, x_1^2 + x_2^2 = 1\}$ is a deformation retraction of Ω and thus $i : A \hookrightarrow \Omega$ induces an isomorphism $i^* : H^*_{dR}(\Omega) \longrightarrow H^*_{dR}(A)$.

Bibliographical remarks. *As at the end of the last section plus the parts on de Rham cohomology in [BT], [J] and [KL].*

2.8 Integration of differential forms on manifolds

Definitions. Let M be a connected manifold of dimension m.

(1) An "orientation form on M" is an element Ω in $\mathcal{E}^m(M)$ such that $\Omega_p \neq 0$ for all p in M, i.e. $((\Omega_p))_{\mathbb{R}} = \Lambda^m(T_pM)^*$ for all p in M.

(2) Let Ω' and Ω'' be two orientation forms on M. Then "Ω' and Ω'' are equivalent (as orientation forms)" if there is a smooth function $f : M \to \mathbb{R}$ such that $f(p) > 0$ for all p in M and $\Omega'' = f \cdot \Omega'$. We write then $[\Omega']_{\text{orientation}} = [\Omega'']_{\text{orientation}}$.

(3) An "orientation on M" is the class $[\Omega]_{\text{orientation}}$ of an orientation form Ω on M.

(4) The manifold M is called "orientable" if there exists an orientation form on M.

(5) An atlas $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ of an oriented manifold $(M, [\Omega]_{\text{orientation}})$ is called "positively oriented (with respect to the orientation $[\Omega]_{\text{orientation}}$)" if for all α in A

$$(\varphi_\alpha^{-1})^*\Omega = g_\alpha \, dx_1^\alpha \wedge \ldots \wedge dx_m^\alpha$$

with a smooth strictly positive function $g_\alpha : V_\alpha = \varphi_\alpha(U_\alpha) \to \mathbb{R}$. (The coordinates on $V_\alpha \subset \mathbb{R}^m$ are denoted by $(x_1^\alpha, \ldots, x_m^\alpha)$.)

Remarks.

(1) An orientation form (respectively an orientation) on a m–dimensional manifold should be intuitively seen as a "smoothy varying orientation form (respectively an orientation) on T_pM for all p in M".

(2) A m–dimensional connected manifold M is orientable if and only if the smooth real line bundle $\Lambda^m T^*M \to M$ is trivializable.

(3) An orientable connected manifold has exactly two orientations.

(4) The notion of a positively oriented atlas is well–defined.

(5) A manifold M is called orientable if each connected component of M is orientable.

(6) Let M and N be manifolds and Ω_M and Ω_N orientation forms on M respectively N. A smooth map $F : M \to N$ is called "orientation–preserving" if $[F^*\Omega_N]_{\text{orientation}} = [\Omega_M]_{\text{orientation}}$.

Examples.

(1) The form $dx_1 \wedge \ldots \wedge dx_m$ on $\mathbb{R}^m$ is called the "canonical orientation form on $\mathbb{R}^m$".

(2) Let $m \geq 2$ and let $f : \mathbb{R}^m \to \mathbb{R}, f(x) = \frac{1}{2}(\|x\|^2 - 1)$ then $\{x \in \mathbb{R}^m \mid f(x) = 0\} = S^m$ and $T_pS^m = \{v \in T_p\mathbb{R}^m \mid (df)_p(v) = 0\}$ for all p in S^{m-1}. Furthermore $U := \mathbb{B}_1(0) = \{x \in \mathbb{R}^m \mid f(x) < 0\}$ and $\partial U = \{x \in \mathbb{R}^m \mid f(x) = 0\} = S^{m-1}$. The "(outward pointing) normal field on S^{m-1}" is defined as

$$N(p) = \left(\sum_{j=1}^{m} \left(\frac{\partial f}{\partial x_j}\right)^2\right)^{-1/2} \cdot \left(\sum_{j=1}^{m} \frac{\partial f}{\partial x_j}\frac{\partial}{\partial x_j}\right)(p).$$

Note that N is *not* a vector field on S^{m-1} but the restriction of a vector field defined on an open neighborhood of S^{m-1} in $\mathbb{R}^m$ to S^{m-1}, i.e. a section of $(T\mathbb{R}^m)|_{S^{m-1}}$.

More concretely we have the formula $N(p) = \sum_{j=1}^m p_j \frac{\partial}{\partial x_j}|_p$. Let furthermore $\Lambda := dx_1 \wedge \ldots \wedge dx_m$ be the canonical orientation on $\mathbb{R}^m$ and let us set $\Omega_p := i_{N(p)}\Lambda_p = (i_N\Lambda)_p|_{T_pS^{m-1}}$. Then Ω is a non–vanishing $(m-1)$–form on S^{m-1}, in explicit terms

$$\Omega_p = (\sum_{j=1}^{m}(-1)^j x_j dx_1 \wedge \ldots \wedge \widehat{dx_j} \wedge \ldots \wedge dx_m)_p.$$

The orientation $[\Omega]_{\text{orientation}}$ on S^{m-1} is called the "canonical induced boundary orientation (with respect to $S^{m-1} = \partial U$ and the orientation $[\Lambda]_{\text{orientation}}$)".

(3) Let again $m \geq 2$ and $\mathbb{P}_{m-1}(\mathbb{R}) = \mathbb{P}(\mathbb{R}^m)$. Denoting the canonical projection $\mathbb{R}^m \backslash \{0\} \to \mathbb{P}(\mathbb{R}^m), x \mapsto [x]$ by $\tilde{\pi}$, we have a smooth surjective map of constant rank $m-1$ defined by

$$\pi : S^{m-1} \to \mathbb{P}_{m-1}(\mathbb{R}), \quad \pi(x) := \tilde{\pi}(x) = [x].$$

For any $(m-1)$–form Θ on $\mathbb{P}_{m-1}(\mathbb{R})$ we have the pullback $\pi^*\Theta$ in $\mathcal{E}^{m-1}(S^{m-1})$. Since $\mathbb{P}_{m-1}(\mathbb{R}) \cong S^m/_\sim$, where $x \sim y$ if and only if either $y = x$ or $y = -x =: \tau(x)$, a differential form η on S^{m-1} is the pullback of a form on $\mathbb{P}_{m-1}(\mathbb{R})$ if and only if $\tau^*\eta = \eta$. Obviously

$\tau^*\Omega = (-1)^m\Omega$ for Ω the orientation form on S^{m-1} constructed in Example (2). It follows that $\mathbb{P}_{m-1}(\mathbb{R})$ is orientable for m an even integer with $m \geq 2$.

Let now be $m \geq 2$ and odd. Assuming that $\mathbb{P}_{m-1}(\mathbb{R})$ is orientable, the pull–back of an orientation form Θ is a multiple of Ω : $\pi^*\Theta = g \cdot \Omega$ with g a smooth function on S^{m-1}. Since π has constant rank equal to $m-1$, $\pi^*\Theta$ is an orientation form on S^{m-1} and, after possibly changing Θ to $(-\Theta)$, we can assume that $g > 0$ on S^{m-1} since Ω is an orientation form. The identity

$$g \cdot \Omega = \tau^*(g \cdot \Omega) = \tau^*(g) \cdot \tau^*(\Omega) = (g \circ \tau)(-\Omega)$$

implies that g is a strictly positive smooth function on S^{m-1} fulfilling $g(-x) = -g(x)$ for all x in S^{m-1}. Since there are no such functions g it follows that $\mathbb{P}_{2n}(\mathbb{R})$ is not orientable for $n \geq 1$.

Definitions. Let M be a manifold and $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ an atlas of M.

(1) The atlas $\mathfrak{A}$ is called "locally finite" if for all p in M the number of α in A such that p is in U_α is finite.

(2) Let $\mathfrak{A}$ be now a locally finite atlas. A collection of smooth functions $\{\chi_\alpha \,|\, \alpha \in A\}$ on M is called a "partion of unity (subordinate to $\mathfrak{A}$)" if the following conditions are satisfied:

(I) the values of χ_α are in $[0,1]$ for all α in A,

(II) The closed set $\operatorname{supp}\chi_\alpha = \overline{\{x \in M | \chi_\alpha(x) \neq 0\}}$ is contained in U_α,

(III) For all p in M one has $\sum_{\alpha \in A} \chi_\alpha(p) = 1$.

(Note that the sum in (III) is finite since $\mathfrak{A}$ is locally finite.)

Proposition. *Let M be a manifold. Then there exists a locally finite atlas and for each locally finite atlas there exists a partition of unity.*

Remark. We will not give a proof of the preceding proposition, but we stress at this point that we included the conditions of Hausdorff and second–countability in our definition of a manifold. These conditions assure the existence of partitions of unity. (See textbooks on manifolds as [AMR] for details.)

As a first application we note the following

Proposition. *Let M be a m–dimensional manifold. Then the following are equivalent:*

(i) M is orientable.

(ii) M has a locally finite atlas that is positively oriented with respect to an orientation form Ω on M.

(iii) M has a locally finite atlas $\{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ such that the Jacobi determinants $\det((\varphi_{\alpha\beta})_*)$ *are everywhere positive for all α, β in A.*

Proof. Let M be oriented by the orientation form Ω and let $\tilde{\mathfrak{A}} = \{(U_\alpha, \tilde{\varphi}_\alpha) \,|\, \alpha \in A\}$ be any locally finite atlas. We define $\varphi_\alpha : U_\alpha \to \mathbb{R}^m$ as follows: if $(\tilde{\varphi}_\alpha^{-1})^*\Omega = \tilde{g}_\alpha \, dx_1^\alpha \wedge \ldots \wedge dx_m^\alpha$ and $\tilde{g}_\alpha$ is everywhere positive, we set $\tilde{\varphi}_\alpha := \varphi_\alpha$. If $\tilde{g}_\alpha$ is everywhere negative then we define $\varphi_\alpha := \Psi \circ \tilde{\varphi}_\alpha$, where $\Psi(x_1^\alpha, x_2^\alpha, \ldots, x_m^\alpha) := (-x_1^\alpha, x_2^\alpha, \ldots, x_m^\alpha)$. It follows that $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ is a positively oriented atlas (with respect to Ω).

Given any atlas $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ and $(\varphi_\alpha^{-1})^*\Omega = g_\alpha dx_1^\alpha \wedge \ldots \wedge dx_m^\alpha$, we calculate

$$(\det((\varphi_{\alpha\beta})_*))\cdot dx_1^\beta \wedge \ldots \wedge dx_m^\beta = \varphi_{\alpha\beta}^*(dx_1^\alpha \wedge \ldots \wedge dx_m^\alpha) = (\varphi_\beta^{-1})^* \circ (\varphi_\alpha)^*(\frac{1}{g_\alpha}\cdot g_\alpha\, dx_1^\alpha \wedge \ldots \wedge dx_m^\alpha)$$

$$= (\varphi_\beta^{-1})^*(\frac{1}{g_\alpha \circ \varphi_\alpha}\Omega) = (\frac{1}{g_\alpha \circ \varphi_\alpha \circ \varphi_\beta^{-1}})(\varphi_\beta^{-1})^*\Omega = \frac{g_\beta}{g_\alpha(\varphi_{\alpha\beta})}\, dx_1^\beta \wedge \ldots dx_m^\beta.$$

Thus for a positively oriented atlas we have $(\det(\varphi_{\alpha\beta})_*) > 0$.

It remains only to prove that (iii) implies (i). Given a locally finite atlas $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ such that $(\det((\varphi_{\alpha\beta})_*)) > 0$ and a partition of unity $\{\chi_\alpha \,|\, \alpha \in A\}$ subordinate to $\mathfrak{A}$ we define differential forms $\Omega_\alpha := \chi_\alpha \cdot \varphi_\alpha^*(dx_1^\alpha \wedge \ldots \wedge dx_m^\alpha)$ on M. Since $\Omega_\alpha(p) = 0$ for all but a finite number of α for each p in M, the differential m–form

$$\Omega := \sum_{\alpha \in A} \Omega_\alpha$$

is well–defined on M. Given p in U_β we have (with $A_p = \{\alpha \in A \,|\, p \in U_\alpha\}$)

$$((\varphi_\beta^{-1})^*\Omega)_{\varphi_\beta(p)} = \sum_{\alpha \in A_p} ((\varphi_\beta^{-1})^*\Omega_\alpha)_{\varphi_\beta(p)} = \sum_{\alpha \in A_p} [(\chi_\alpha \circ \varphi_\beta^{-1})\cdot((\varphi_\beta^{-1})^* \circ \varphi_\alpha^*(dx_1^\alpha \wedge \ldots \wedge dx_m^\alpha))]_{\varphi_\beta(p)}$$

$$= \sum_{\alpha \in A_p} \chi_\alpha(p) \cdot (\varphi_{\alpha\beta}^*(dx_1^\beta \wedge \ldots \wedge dx_m^\beta))_{\varphi_\beta(p)} = \Big[\sum_{\alpha \in A_p} \chi_\alpha(p) \cdot (\det((\varphi_{\alpha\beta})_{*p}))\Big] \cdot dx_1^\beta \wedge \ldots \wedge dx_m^\beta$$

and thus Ω is a nowhere–vanishing m–form on M, i.e. an orientation form (and $\mathfrak{A}$ a positively oriented atlas with respect to Ω). □

Corollary. *Let M be a connected manifold with $m = \dim_{\mathbb{R}} M \geq 2$ and U an open subset of M such that ∂U is a closed $(m-1)$–dimensional submanifold of M and such that $\partial U \cap \mathring{\bar{U}}$ ($\mathring{\bar{U}}$ denotes the interior of the closure of U) is empty. If M is oriented then ∂U has a "canonical induced boundary orientation".*

Proof. Let $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ be a locally finite oriented atlas of M such that for $N := \partial U$ we have $\varphi_\alpha(U_\alpha \cap N) = \{(x_1^\alpha, \ldots, x_m^\alpha) \in V_\alpha = \varphi_\alpha(U_\alpha) \,|\, x_1^\alpha = 0\}$ and $\varphi_\alpha(U_\alpha \cap U) = \{(x_1^\alpha, \ldots, x_m^\alpha) \in V_\alpha \,|\, x_1^\alpha < 0\}$. Let $B = \{\alpha \in A \,|\, U_\alpha \cap N \neq \emptyset\}$ and $\Psi_\alpha : U_\alpha \cap N \to \mathbb{R}^{m-1}$ be defined by $\Psi_\alpha(p) := (y_1^\alpha(p), \ldots, y_{m-1}^\alpha(p)) := (x_2^\alpha(p), \ldots, x_m^\alpha(p))$. Then $\mathcal{B} = \{(U_\alpha \cap N, \Psi_\alpha) \,|\, \alpha \in B\}$ is a locally finite atlas for N and it is easy to check that $\det((\Psi_{\alpha\beta})_*) > 0$ for all α, β in B. By the preceding proposition there exists an orientation form on $N = \partial U$ such that $\mathcal{B}$ is positively oriented with respect to this "canonical boundary orientation".□

Exercise. Show that the boundary orientation we constructed on S^{m-1} is a special case of the preceding corollary.

Definition. Let V be open in $\mathbb{R}^m$ and let $\Lambda = g \cdot dx_1 \wedge \ldots \wedge dx_m$ be in $\mathcal{E}^m(V)$ such that $\operatorname{supp} g$ is compact. Then

$$\int_V \Lambda := \int_V g\, dx_1 \ldots dx_m,$$

where the right–hand side is defined by iterated integration (in the sense of the Lebesgue or the Riemann integral).

Lemma. *Let V' and V'' be open sets in $\mathbb{R}^m$, oriented by the standard orientation form of $\mathbb{R}^m$. Let furthermore Λ be in $\mathcal{E}^m(V'')$ with compact support, and $\varphi : V' \to V''$ an orientation–preserving diffeomorphism. Then*

$$\int_{V'} \varphi^*(\Lambda) = \int_{V''} \Lambda.$$

Proof. Using the transformation formula for multiple integrals we find

$$\int_{\varphi(V')=V''} \Lambda = \int_{\varphi(V')} g(y)dy_1 \ldots, dy_m = \int_{V'} ((\det \varphi_*)(x)) \cdot g(\varphi(x)) \cdot dx_1 \ldots dx_m$$

$$= \int_{V'} ((\det \varphi_*)(x)) \cdot g(\varphi(x))\, dx_1 \wedge \ldots \wedge dx_m = \int_{V'} g(\varphi(x)) d\varphi_1 \wedge \ldots \wedge d\varphi_m = \int_{V'} \varphi^*\Lambda.$$

□

Definition. Let M be an oriented m–dimensional manifold and Λ a m–form with compact support. Then

$$\int_M \Lambda := \sum_{\alpha\in A} \int_{\varphi_\alpha(U_\alpha)} (\varphi_\alpha^{-1})^*(\chi_\alpha \cdot \Lambda),$$

where $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ is a locally finite, positively oriented atlas for M and $\{\chi_\alpha \,|\, \alpha \in A\}$ is a partition of unity subordinate to $\mathfrak{A}$.

Exercise. Show that $\int_M \Lambda$ is well–defined, i.e. independent of the chosen locally finite, positively oriented atlas and the chosen subordinate partition of unity. (Hint: Use the preceding lemma.)

Theorem ("Stokes' theorem"). *Let M be a connected, oriented, m–dimensional manifold with $m \geq 2$, and η in $\mathcal{E}^{m-1}(M)$. Let furthermore U be an open subset of M such that its closure $\bar{U}$ is compact and its boundary ∂U is a smooth closed submanifold of M. Then*

$$\int_U d\eta = \int_{\partial U} \eta.$$

Proof. The detailed derivation of this theorem can be found in many textbooks on manifolds. For a short proof see, e.g., [BT], pp. 31. □

Remarks.

(0) A purist whould introduce the injection $j : \partial U \to M, j(p) = p$ and write Stokes' formula as follows:

$$\int_U d\eta = \int_{\partial U} j^*(\eta).$$

It obviously follows that the integral of $j^*(\eta)$ over boundary components of dimension strictly smaller than $m-1$ vanishes since they have no non–zero $(m-1)$–forms.

(1) Given our preparations the proof of Stokes' theorem is reduced to an ingenious reduction to the "Fundamental theorem of calculus": if $F : \mathbb{R} \to \mathbb{R}$ smooth and $a < b$ then $F(b) - F(a) = \int_a^b F'(x)\,dx$.

Though formally not included in the above formulation of Stokes' theorem it fits in the following sense:

Let $M = \mathbb{R}$, $U = (a, b)$ and $\eta = F$ in $\mathcal{E}^0(M) = \mathcal{E}^{m-1}(M)$. The outward pointing normal vectors in $\partial U = \{a, b\}$ are then

$$N(a) = -\frac{d}{dx}\Big|_a \quad \text{and} \quad N(b) = \frac{d}{dx}\Big|_b,$$

i.e. parallel to the positively oriented basis $\frac{d}{dx}$ in b and anti–parallel in a. Thus "the integral of the zero–form F over ∂U with respect to the boundary orientation" should be $(-F(a))+F(b)$, i.e.

$$F(b) - F(a) = \int_{\partial U} F = \int_U dF = \int_U F'(x)\,dx = \int_a^b F'(x)\,dx.$$

(2) The usual integral theorems known from vector calculus in $\mathbb{R}^2$ and $\mathbb{R}^3$ are special cases of Stokes' theorem. As an example we will give the following.

Corollary (Gauss' theorem). *Let V be open in $\mathbb{R}^3$ and $\vec{K} = {}^t(K_1, K_2, K_3) : V \to \mathbb{R}^3$ be a smooth force field. Let furthermore be U open in V such that its closure $\bar{U}$ is compact and contained in V and such that ∂U is a smooth closed submanifold of V. Then we have for $\eta = i_K\Lambda$ (with $K = \sum_{j=1}^3 K_j \frac{\partial}{\partial x_j}$ and $\Lambda = dx_1 \wedge dx_2 \wedge dx_3$):*

$$\int_U d(i_K\Lambda) = \int_{\partial U} i_K\Lambda.$$

Proof. Obviously the assertion of the corollary is a special case of Stokes' theorem. □

Remark. The interesting part of the corollary is given by a further translation into vector calculus. First, a direct calculation shows that $d(i_K\Lambda) = \left(\sum_{j=1}^3 \frac{\partial K_j}{\partial x_j}\right) \cdot \Lambda = (\operatorname{div}(K)) \cdot \Lambda = (\vec{\nabla} \cdot \vec{K}) \cdot \Lambda$. Secondly, in a point p of ∂U the two–form $i_K\Lambda$ restricted to $T_p\partial U$ is necessarily proportional to the canonical orientation $(i_N\Lambda)_p = i_{N(p)}\Lambda_p$, where $N(p)$ in T_pV is uniquely fixed as the outward pointing normal vector such that $\|N(p)\| = 1$, $N(p)$ is orthogonal to $T_p\partial U \subset T_pV = T_p\mathbb{R}^3$ and $p + \epsilon\vec{N}(p)$ is outside U for small $\epsilon > 0$. (Here $N(p) = \sum_{j=1}^3 N_j(p)\frac{\partial}{\partial x_j}|_p$ and $\vec{N} = {}^t(N_1, N_2, N_3)$ of course.)

Let us remark that $N(p) = \left(\sum_{j=1}^3 |\frac{\partial g}{\partial x_j}(p)|^2\right)^{-\frac{1}{2}} \cdot \left(\sum_{j=1}^3 \frac{\partial g}{\partial x_j}(p) \cdot \frac{\partial}{\partial x_j}|_p\right)$ if g is a local function near p such that $\{g < 0\} = U$ and $\{g = 0\} = \partial U$ and $dg|_{\partial U} \neq 0$ as considered before in this section.

We fix an ordered orthonormal basis $\{v_1, v_2\}$ of $T_p\partial U$ such that $(i_N\Lambda)_p(v_1, v_2) > 0$ (and then equal to one in fact) and calculate:

$$(i_{K(p)}\Lambda_p)(v_1, v_2) = \Lambda_p(K(p), v_1, v_2) = \Lambda_p((\vec{K}(p) \cdot \vec{N}(p))\vec{N}(p), v_1, v_2),$$

since $\{N(p), v_1, v_2\}$ is an orthonormal basis of $T_pV = T_p\mathbb{R}^3$ and thus $(i_K\Omega)_p(v_1, v_2) = ((\vec{K} \cdot \vec{N})(p))(i_N\Omega)_p(v_1, v_2)$, i.e. $i_K\Omega = (\vec{K} \cdot \vec{N})i_N\Omega$. Interpreting $\vec{N}(i_N\Omega)$ as the "vectorial surface element $d\vec{S}$" we arrive at a formulation of Gauss' theorem which is frequently found in the physics literature:

$$\int_U (\vec{\nabla} \cdot \vec{K}) dx_1\, dx_2\, dx_3 = \int_{\partial U} \vec{K} \cdot d\vec{S}.$$

Proposition. *Let M be a manifold and N a compact submanifold with an orientation. Then the integral over N defines a linear functional on $H^*_{dR}(M)$.*

Proof. For μ in $\mathcal{E}^k(M)$ let $\int_N \mu$ denote the value of $\int_N i_N^*(\mu)$ defined by the orientation of N. (The inclusion map $N \hookrightarrow M$ is denoted by i_N here.)

Let us assume that $\mu = d\eta$ for η in $\mathcal{E}^{k-1}(M)$. Then $i_N^*(\mu) = d(i_N^*(\eta))$ and with Stokes' theorem we will show that $\int_N i_N^*(\mu) = 0$.

Let p be any point in N and $\varphi : W \to V$ be a chart of N such that $\varphi(p) = 0$ and such that $\overline{\mathbb{B}_{\epsilon_0}(0)} \subset V$ for a $\epsilon_0 > 0$. Let $U_\epsilon := N \setminus \varphi^{-1}(\overline{\mathbb{B}_\epsilon(0)})$ for $0 < \epsilon < \epsilon_0$ then by Stokes' theorem we have

$$\int_N d(i_N^*\eta) = \lim_{\epsilon \to 0} \int_{U_\epsilon} d(i_N^*\eta) = \lim_{\epsilon \to 0} \int_{\partial U_\epsilon} i_N^*\eta.$$

On the other hand

$$\int_{\partial U_\epsilon} i_N^*\eta = \int_{||x||=\epsilon} (\varphi^{-1})^*(i_N^*\eta)$$

and thus converges for $\epsilon \searrow 0$ to zero.

It follows that $\int_N d\eta = 0$ and thus the map

$$H^k_{dR}(M) \to \mathbb{R}, \quad [\mu] \mapsto \int_N \mu$$

is a well–defined linear functional. □

Corollary. *Let M be a compact connected orientable m–dimensional manifold. Then $[\Omega] \neq 0$ in $H^m_{dR}(M)$ for all nowhere–vanishing m–forms Ω on M.*

Proof. Let Ω_0 be a m–form defining an orientation of M. Then $\Omega = f \cdot \Omega_0$ for a nowhere–vanishing smooth function f on M. We may assume without loss of generality that $f > 0$.

Let $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ be any locally finite, positively oriented atlas of M and $\{\chi_\alpha \,|\, \alpha \in A\}$ any partition of unity subordinate to $\mathfrak{A}$. Then

$$\int_M \Omega = \int_M f \cdot \Omega_0 = \sum_\alpha \int_{\varphi_\alpha(U_\alpha)} (\varphi_\alpha^{-1})^*(\chi_\alpha f \Omega_0) = \sum_\alpha \Big[\int_{\varphi_\alpha(U_\alpha)} ((\chi_\alpha f) \circ \varphi_\alpha^{-1}) \cdot (\varphi_\alpha^{-1})^*(\Omega_0) \Big]$$

and all summands on the last right hand side are non–negative, and at least one is strictly positive. Thus $\int_M \Omega > 0$.

Since the functional

$$H^m_{dR}(M) \to \mathbb{R}, \quad [\mu] \mapsto \int_M \mu$$

is well–defined by the preceding proposition it follows that Ω cannot be exact, i.e. $[\Omega] \neq 0$ in $H^m_{dR}(M)$. □

Remark. With the hitherto developped theory it is possible to sharpen the assertion of the last proposition to the statement that integration over M is, in the compact case, an isomorphism from $H^m_{dR}(M)$ to $\mathbb{R}$. (See [AMR], pp. 552.)

Bibliographical remarks. *As for Section 2.7.*

3. Symplectic geometry

3.1 Symplectic manifolds

Definitions. Let M be a (real) manifold.

(1) A differential 2–form ω on M is called an "almost–symplectic form" if and only if

$$\ker \omega_p = \{0\} \quad \text{for all} \quad p \text{ in } M,$$

i.e. "ω is everywhere non–degenerate". A pair (M, ω) consisting of a manifold M and an almost–symplectic form ω is called an "almost–symplectic manifold".

(2) An almost–symplectic form ω on M is called a "symplectic form" if and only if ω is closed, i.e. $d\omega = 0$. A pair (M, ω) consisting of a manifold M and a symplectic form ω is called a "symplectic manifold".

Lemma. *Let M be a manifold and ω in $\mathcal{E}^2(M)$. Then the following are equivalent:*

(i) ω is an almost–symplectic form

(ii) $\omega_p^\flat : T_pM \to (T_pM)^* = (T^*M)_p$ is an isomorphism for all p in M.

If M is furthermore of pure dimension m then (i) and (ii) are equivalent to

(iii) $\omega^{[m/2]}$ is an orientation form on M.

Proof. The assertions follow directly from Section 1.3 and the definition of an orientation form. □

Corollary. *Let (M, ω) be an almost–symplectic manifold of dimension m. Then M is orientable and m is even.*

Proof. The corollary follows immediately from the preceding lemma. □

Definition. Let (M,ω) be a symplectic manifold of dimension $m=2n$. The orientation form

$$\Omega := \left((-1)^{\frac{(n-1)n}{2}} \cdot \frac{1}{n!}\right)\omega^n$$

is called the "canonical orientation form (or Liouville form) on (M,ω)". The associated orientation $[\Omega]_{\text{orientation}}$ is called the "canonical orientation on (M,ω)".

Proposition. *Let M be a connected compact manifold of dimension m. Let furthermore ω be a symplectic form on M and $c=[\omega]$ in $H^2_{dR}(M)$ be the de Rham cohomology class of ω. Then the de Rham cohomology classes c^k are non–zero in $H^*_{dR}(M)$ for $k=1,2,\dots,\frac{m}{2}$.*

Proof. Since ω is closed it defines a de Rham cohomology class $c=[\omega]$.

Assume that there is a k in $\{1,2,\dots,\frac{m}{2}\}$ such that $c^k=0$, i.e. there exists μ in $\mathcal{E}^{2k-1}(M)$ such that $d\mu=\omega^k$. Closedness of ω implies that $\omega^{\frac{m}{2}}=d\eta$ with $\eta:=\mu\wedge\omega^{(\frac{m}{2}-k)}$. Since $\omega^{\frac{m}{2}}$ is nowhere vanishing $\int_M \omega^{\frac{m}{2}}\neq 0$, where the integral is defined by one of the two possible orientations of M, e.g. by the canonical orientation of (M,ω). Since it was shown in Section 2.8 that $\int_M d\eta=0$ for all η in $\mathcal{E}^{m-1}(M)$ we arrive at a contradiction. Thus there is no k in $\{1,\dots,\frac{m}{2}\}$ such that ω^k is exact. $\square$

Remark. The preceding results of this section show that not all manifolds can carry a symplectic form.

Examples.

(1) $M=\mathbb{R}^{2n}$ and $\omega:=\sum_{j=1}^n dx_j\wedge dx_{n+j}$. Then (M,ω) is symplectic and

$$\Omega = \left((-1)^{\frac{(n-1)n}{2}}\cdot\frac{1}{n!}\right)\omega^n = dx_1\wedge dx_2\wedge\ldots\wedge dx_{2n-1}\wedge dx_n.$$

The form ω is sometimes called the "standard symplectic form on $\mathbb{R}^{2n}$."

(2) Let Σ be a two–dimensional orientable manifold and Ω an orientation form on Σ. Then $\omega:=\Omega$ is a symplectic form on Σ (and $\Omega=\omega$ is the canonical orientation form on (Σ,ω)).

(2.1) Let $\Sigma=S^2$. In the notations of Section 2.8 there is a natural orientation form $\Omega=(i_N\Lambda)|_{TS^2}$ (with $\Lambda=dx_1\wedge dx_2\wedge dx_3$ and $N=\sum_{j=1}^3 x_j\frac{\partial}{\partial x_j}$). Then (S^2,Ω) is a symplectic manifold.

(2.2) Let $M=S^1\times S^1$ and $\Omega=d\vartheta_1\wedge d\vartheta_2$, where $d\vartheta_j(\frac{\partial}{\partial\vartheta_k})=\delta_{j,k}$. Then $(S^1\times S^1,\Omega)$ is a symplectic manifold.

(3) Let (M_1,ω_1) and (M_2,ω_2) be symplectic manifolds and λ_1,λ_2 in $\mathbb{R}\backslash\{0\}$. Then the form $\lambda_1\cdot(\mathrm{pr}_1^*\omega_1)+\lambda_2\cdot(\mathrm{pr}_2^*\omega_2)$ is a symplectic form on $M_1\times M_2$.

(4) Let (M,ω) be a symplectic manifold and U be open in M. Then $(U,\omega|_U)$ is a symplectic manifold.

Proposition. *Let Q be a manifold, $M:=T^*Q$ its cotangent bundle and $\pi=\pi^{T^*Q}:T^*Q\to Q$ the canonical projection. Then M has a canonical differential one–form $\theta:=\theta^{T^*Q}$ defined by*

$$\theta_{\alpha_q}(v_{\alpha_q}) := \alpha_q((\pi_*)_{\alpha_q}(v_{\alpha_q})) \textit{ for all } q \textit{ in } Q, \alpha_q \textit{ in } (T^*Q)_q=(T_qQ)^* \textit{ and } v_{\alpha_q} \textit{ in } T_{\alpha_q}(T^*Q).$$

Furthermore, the two–form $\omega := \omega^{T^*Q} := -d\theta$ is a symplectic form on $M = T^*Q$.

Remark. The form θ (respectively ω) on T^*Q is often called the "canonical one–form (respectively two–form) on the cotangent bundle T^*Q".

Proof of the proposition. Let $\varphi : U \to V \subset \mathbb{R}^n$ be a coordinate chart with domain U, an open set in Q. Then there is an induced vector bundle isomorphism $\Phi = (T\varphi)^* : T^*V \to T^*U$ over $\varphi^{-1} : V \to U$ defined as follows:

$$\Phi|_{T_x^*V} = (T_{\varphi^{-1}(x)}\varphi)^* : T_x^*V = (T_xV)^* \to (T_{\varphi^{-1}(x)}U)^* = T^*_{\varphi^{-1}(x)}U \quad \text{for all } x \text{ in } V.$$

Furthermore we trivialize T^*V as usual: let $x = {}^t(x_1, \dots, x_n)$ in $V \subset \mathbb{R}^n$ and $y = {}^t(y_1, \dots, y_n)$ in $\mathbb{R}^n$, then we set

$$\Psi : V \times \mathbb{R}^n \to T^*V, \psi(x,y) := \sum_{j=1}^n y_j \, dx_j|_x.$$

Clearly Ψ is a vector bundle isomorphism over Id_V.

We describe elements of $T_{(x,y)}(V \times \mathbb{R}^n)$ by pairs $(\xi, \eta) = \Big(\sum_{j=1}^n \xi_j \frac{\partial}{\partial x_j}|_x, \ \sum_{j=1}^n \eta_j \frac{\partial}{\partial y_j}|_y\Big)$ with $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_n$ in $\mathbb{R}$ and calculate:

$$((\Phi \circ \Psi)^*\theta)_{(x,y)}(\xi,\eta) = \theta_{\Phi(\Psi(x,y))}((\Phi_* \circ \Psi_*)(\xi,\eta)) = \Phi(\Psi(x,y))(((\pi^{T^*Q})_* \circ \Phi_* \circ \Psi_*)(\xi,\eta)).$$

Since $(\pi^{T^*M} \circ \Phi \circ \Psi)(x,y) = \varphi^{-1}(x)$ it follows

$$\begin{aligned}((\Phi \circ \Psi)^*\theta)_{(x,y)}(\xi,\eta) &= ((T\varphi)^*(\Psi(x,y)))((T\varphi^{-1})(\xi)) = (\Psi(x,y))(\xi) \\ &= \Big(\sum_{j=1}^n y_j \, dx_j|_x\Big)\Big(\sum_{j=1}^n \xi_j \frac{\partial}{\partial x_j}|_x\Big) = \sum_{j=1}^n y_j\xi_j = \Big(\sum_{j=1}^n y_j \, dx_j\Big)_{(x,y)}(\xi,\eta),\end{aligned}$$

i.e. $(\Phi \circ \Psi)^*\theta = \sum_{j=1}^n y_j \, dx_j$ in $\mathcal{E}^1(V \times \mathbb{R}^n)$.

It follows from this local computation that θ and $\omega = -d\theta$ are (smooth) differential forms on $M = T^*Q$. Furthermore we have

$$(\Phi \circ \Psi)^*\omega = (\Phi \circ \Psi)^*(-d\theta) = -d((\Phi \circ \Psi)^*\theta) = \sum_{j=1}^n dx_j \wedge dy_j$$

and hence ω is non–degenerate in all points of the open set T^*U in T^*Q. Since the above local calculations are valid for all charts of Q and since the cotangent bundles over chart domains of Q form an admissible atlas of T^*Q it follows that ω is everywhere non–degenerate on $M = T^*Q$. Closedness of ω is trivial since $d \circ d = 0$ and thus ω is a symplectic form on T^*Q.

□

Remark. If we describe elements of $V \times \mathbb{R}^n$ (V open in $\mathbb{R}^n$) by $(q,p) = ({}^t(q_1, \dots, q_n), {}^t(p_1, \dots, p_n))$ the preceding proof yields the traditional formulas:

$$(\Phi \circ \Psi)^*\theta = \sum_{j=1}^n p_j dq_j \quad \text{and} \quad (\Phi \circ \Psi)^*\omega = \sum_{j=1}^n dq_j \wedge dp_j$$

for the cotangent bundle forms θ and ω.

Exercise. Show that the canonical cotangent bundle two–form $\omega^{T^*\mathbb{R}^n}$ on $T^*\mathbb{R}^n$ is pulled–back by $\Phi \circ \Psi = \Psi$ to the standard symplectic two–form on $\mathbb{R}^{2n}$ (up to renaming of the variables $y_j =: x_{n+j}$ for $j = 1, \ldots, n$).

Remark. Given a manifold Q, a "configuration space", the symplectic manifold (T^*Q, ω) with the cotangent bundle two–form ω is often called the "phase space (associated to Q)". Furthermore the dimension of Q is often referred to as the "number of degrees of freedom".

Bibliographical remarks. *The classics for the mathematical treatment of mechanics in the symplectic language are of course [AM] and [Ar1]. We would like to add [Be], [Bry] and [GS].*

3.2 Maps and submanifolds of symplectic manifolds

Definitions. Let (M, ω_M) and (N, ω_N) be symplectic manifolds and $F : M \to N$ a smooth map.

(1) The map F is called "symplectic" if $F^*(\omega_N) = \omega_M$.

(2) If F is a diffeomorphism such that $F^*(\omega_N) = \omega_M$, F is called a "symplectic diffeomorphism" (or sometimes also a "symplectomorphism").

Proposition. *Let (M, ω_M) and (N, ω_N) be pure–dimensional symplectic manifolds and $F : M \to N$ a symplectic map. Then*

(i) $\quad T_pF = (F_*)_p : (T_pM, (\omega_M)_p) \to (T_{F(p)}N, (\omega_N)_{F(p)})$

is a symplectic linear map, i.e. $(T_pF)^*((\omega_N)_{F(p)}) = (\omega_M)_p$ *for all p in M.*

Thus T_pF is injective for all p in M, so that in particular $\dim_{\mathbb{R}} M \leq \dim_{\mathbb{R}} N$ *and the rank of F in p equals the dimension of M (for all p in M).*

If furthermore $\dim_{\mathbb{R}} M = \dim_{\mathbb{R}} N =: 2n$ *then*

(ii) F is a local diffeomorphism, the local inverses are also symplectic, and

(iii) $F^*(\Omega_N) = \Omega_M$, i.e. F preserves the canonical orientation forms and hence, a fortiori, the canonical orientations. If furthermore F is a diffeomorphism then its inverse is also symplectic.

Proof. Exercise using Section 1.4. $\square$

Examples.

(1) Let $f : Q \to Q$ be a diffeomorphism of a manifold Q and $F := (Tf^{-1})^* : T^*Q \to T^*Q$, defined by $F(\alpha_q) = (T_{f(q)}f^{-1})^*(\alpha_q)$ for all q in Q. Then F is a vector bundle isomorphism

over f and $F^*(\theta^{T^*Q}) = \theta^{T^*Q}$. Thus we have of course $F^*(\omega^{T^*Q}) = \omega^{T^*Q}$, i.e., F is a symplectic diffeomorphism of T^*Q. Since F is induced from a diffeomorphism of the configuration space Q, it is traditionally called a "point transformation (of T^*Q)".

(2) Let Σ be an orientable two–dimensional manifold and Ω an orientation form on Σ. Then a diffeomorphism of Σ (with the symplectic form $\omega := \Omega$) is obviously symplectic if and only if it is volume–preserving.

(3) Let g be in $GL(m, \mathbb{R}) \subset \mathrm{Mat}(m \times m, \mathbb{R})$. Then $T_g : \mathbb{R}^m \to \mathbb{R}^m$, $T_g(x) = g \cdot x$ is a diffeomorphism of $\mathbb{R}^m$ with

$$((T_g)_*)_x \Big(\sum_{k=1}^{m} v_k \frac{\partial}{\partial x_k}\Big|_x \Big) = \sum_{j,k=1}^{m} (g_{jk} v_k) \frac{\partial}{\partial x_j}\Big|_{g(x)},$$

i.e., $((T_g)_*)_x$ can be identified with the (linear!) map T_g. Furthermore we define for A in $\mathrm{Mat}(m \times m, \mathbb{R})$ a vector field X_A by $X_A(x) := \sum_{j,k=1}^{m} a_{jk} x_k \frac{\partial}{\partial x_j}|_x$. Then the flow $\varphi^A := \varphi^{X_A}$ of such a "linear vector field on $\mathbb{R}^m$" is easily seen to be given by

$$\varphi^A(t, x) = (e^{tA}) \cdot x = T_{e^{tA}}(x).$$

Taking $m = 2n$ and $\omega_0 := \sum_{j=1}^{n} dx_j \wedge dx_{n+j}$ it follows that the diffeomorphism T_g is symplectic if and only if g is in $Sp(2n, \mathbb{R})$,

In particular for $n = 1$ we find that φ_t^A is symplectic if and only if $\mathrm{trace}(A) = 0$.

Exercise. Fill in the details in Examples (1) and (3).

Definition. Let G be a Lie group and (M, ω) a symplectic manifold. Then a (smooth left–)action $\vartheta : G \times M \to M$ is called a "symplectic action" if $\vartheta_g^*(\omega) = \omega$ for all g in M.

Examples.

(1) If Q is a manifold, G a Lie group and $\vartheta : G \times Q \to Q$ an action, then for all g in G the map $\widehat{\vartheta}_g := (\vartheta_{g-1})^* : T^*Q \to T^*Q$ is a vector bundle isomorphism over ϑ_g such that $(\widehat{\vartheta}_g)^*(\theta^{T^*Q}) = \theta^{T^*Q}$. Thus each $\widehat{\vartheta}_g$ is a symplectic diffeomorphism and one easily sees that $\widehat{\vartheta} : G \times T^*Q \to T^*Q$, $\widehat{\vartheta}(g, \alpha) := \widehat{\vartheta}_g(\alpha)$ is a symplectic action.

(2) The map $\vartheta :\ \mathrm{Sp}(2n, \mathbb{R}) \times \mathbb{R}^{2n} \to \mathbb{R}^{2n}, \vartheta(g, x) := g \cdot x$ is a symplectic action if we supply $\mathbb{R}^{2n}$ with the canonical symplectic form ω_0.

(3) Let $(\mathbb{R}^{2n}, +)$ act on itself by vector addition, i.e., $G := \mathbb{R}^{2n}, M := \mathbb{R}^{2n}, \vartheta : G \times M \to M, \vartheta(a, x) := x + a$. Then ϑ is symplectic, again with respect to the canonical symplectic form on $\mathbb{R}^{2n}$.

(4) Let $\Omega = i_N \Lambda$ be the usual orientation form on the two–sphere $S^2 \subset \mathbb{R}^3$ and let $\vartheta : SO(3, \mathbb{R}) \times S^2 \to S^2$, be defined by $\vartheta(g, x) = g \cdot x$. Then ϑ is a symplectic action with respect to the symplectic form $\omega := \Omega$ on S^2.

Definition. Let M be a $(2n)$–dimensional manifold with symplectic form ω and N a closed k–dimensional submanifold of M. Then

$$(TN)^{\angle} := \{v_p \in T_pM \mid p \in N \quad \text{and} \quad \omega_p(v_p, w_p) = 0 \ \forall\ w_p \in T_pN \subset T_pM\}$$

is called the "skew–complement (or ω–complement) of TN in $TM|_N$".

Lemma. *Let M be a manifold and N a closed submanifold of M. Then*

(i) $TM|_N := \dot{\bigcup}_{p\in N} T_n M$ carries a natural vector bundle structure and TN is a subbundle of $TM|_N$.

If furthermore M carries a symplectic form ω then

(ii) $(TN)^\angle$ is a subbundle of $TM|_N$ and $\dim_\mathbb{R} T_pN + \dim_\mathbb{R}((TN)^\angle)_p = \dim_\mathbb{R} T_pM$ for all p in N.

Proof. *Ad(i).* Given a point p in N there is a chart $\varphi = (x_1, \dots, x_m) : U \to V = \varphi(U) \subset \mathbb{R}^m$ with U open in M and such that $\varphi(U \cap N) = V \cap \{x_1 = \dots = x_k = 0\}$. Working in this chart TM is trivialized by the sections $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_m}$ (over V) and hence so is $TM|_N$ over $\varphi(U \cap N)$. The bundle TN is, again over $\varphi(U \cap N)$ trivialized by $\frac{\partial}{\partial x_{k+1}}, \dots, \frac{\partial}{\partial x_m}$ and thus the first assertion is proven.

Ad(ii). Still working in a chart as above we have now $m = 2n$ for n in $\mathbb{N}$ and we set $\tilde{\omega} = (\varphi^{-1})^*\omega$. Clearly $\left\{ \frac{\partial}{\partial x_{k+1}}\Big|_x, \dots, \frac{\partial}{\partial x_{2n}}\Big|_x \right\}$ is a basis for $T_x\varphi(U \cap N)$ for all x in $\varphi(U \cap N)$ and thus

$$(\varphi_*)_{\varphi^{-1}(x)}((T_{\varphi^{-1}(x)}N)^\angle) = (T_x\varphi(U \cap N))^\angle = \bigcap_{j=k+1}^{2n} \ker\left(\tilde{\omega}_x^\flat\left(\frac{\partial}{\partial x_j}\Big|_x\right)\right).$$

Since $\tilde{\omega}$ is symplectic, the functionals $\tilde{\omega}_x^\flat\left(\frac{\partial}{\partial x_j}\Big|_x\right)$ are linearly independent so that $(TN)^\angle$ is a subbundle of $TM|_N$ and the asserted dimensional formula holds true. $\square$

Remark. Given M and N as in the first part of the preceding proposition, the bundle $TM|_N$ is often called the "tangent bundle of M restricted to N" and with the canonical inclusion $j_N : N \hookrightarrow M, j_N(p) = p$ we can describe it (isomorphically) by the pull–back bundle $(j_N)^*TM$.

Definition. A closed submanifold N of a symplectic manifold (M, ω) is called

(1) "symplectic" if $(TN)^\angle \cap TN = \mathrm{Im}\,(\sigma_0)$, the image of the zero–section $\sigma_0 : N \to TM|_N$,

(2) "isotropic" if $TN \subset (TN)^\angle$,

(3) "coisotropic" if $(TN)^\angle \subset TN$,

(4) "Lagrangian" or a "Lagrange submanifold" if $(TN)^\angle = TN$.

Remark. The first condition is often written $(TN)^\angle \cap TN = \{0\}$.

Lemma. *Let N be a closed submanifold of a symplectic manifold (M, ω) and $j_N : N \hookrightarrow M$ the canonical inclusion of N in M. Then*

(i) N is symplectic if and only if $\omega_p|_{T_pN}$ is non–degenerate for all p in N, i.e. if $(j_N)^*\omega$ is a non–degenerate two–form on N,

(ii) N is isotropic if and only if $\omega_p|_{T_pN}$ is the zero–form for all p in N, i.e. $(j_N)^*\omega = 0$,

(iii) N is Lagrangian if and only if N is isotropic and N has half the dimension of M, i.e. $\dim_{\mathbb{R}} T_pM = 2 \cdot (\dim_{\mathbb{R}} T_pN)$ *for all p in N.*

Proof. Exercise. $\square$

Examples.

(1) Let N be a closed one–dimensional submanifold of a symplectic manifold (M, ω). Then N is isotropic.

(2) Let N be a closed hypersurface of a symplectic manifold (M, ω). Then N is coisotropic.

Proposition. *Let Q be a manifold and η a one–form on Q. Then the image of η, $\eta(Q) = \{\eta(q) \in T_q^*Q \,|\, q \in Q\}$ is a closed submanifold of T^*Q and the map $\eta : Q \to T^*Q$ a diffeomorphism of Q onto $\eta(Q)$.*

Furthermore $\eta(Q)$ is Lagrangian in the symplectic manifold (T^*Q, ω^{T^*Q}) if and only if η is closed.

Proof. The first part follows from the general fact that the image of a section σ of a vector bundle $E \xrightarrow{\pi} M$ is always a closed submanifold such that $\sigma : M \to \sigma(M) \subset E$ is a diffeomorphism with inverse $\pi|_{\sigma(M)}$. (This fact is easily proven by using a local trivialization of $E \xrightarrow{\pi} M$.)

Denoting $\eta(Q)$ by N and the canonical inclusion $N \hookrightarrow T^*Q$ by j_N, the equation $j_N = \eta \circ \pi|_{\eta(Q)}$ implies that it is enough to show that $\eta^*(\omega^{T^*Q}) = 0$ if and only if η is closed, since $\pi|_{\eta(Q)} = N \to Q$ is a diffeomorphism and clearly $\eta(Q)$ has half the dimension of T^*Q.

Let thus q be in Q and u in T_qQ, then

$$(\eta^*(\theta^{T^*Q}))_q(u) = (\theta^{T^*Q})_{\eta(q)}(\eta_*(u)) = \eta(q)((\pi_* \circ \eta_*)(u)) = \eta_q(u),$$

i.e. $\eta^*(\theta^{T^*Q}) = \eta$. It follows that $\eta^*(\omega^{T^*Q}) = -d\eta$ and thus the second part of the proposition is proven. $\square$

Corollary. *Let Q be a manifold and f be a smooth function on Q. Then $(df)(Q) \subset T^*Q$ is Lagrangian with respect to the canonical symplectic form ω^{T^*Q}.*

Proof. Obvious, since $d(df) = 0$. $\square$

Proposition. *Let (M, ω_M) and (N, ω_N) be symplectic manifolds and $F : M \to N$ a diffeomorphism. Then the graph $\Gamma_F := \{(x, y) \in M \times N \,|\, y = F(x)\}$ is a Lagrangian submanifold of the symplectic manifold $(M \times N, (pr_M)^*\omega_M - (pr_N)^*\omega_N)$ if and only if F is symplectic.*

Proof. Let us first observe that for any smooth map $F : M \to N$ between two manifolds the graph Γ_F is a closed submanifold of $M \times N$ and $\widehat{F} : M \to \Gamma_F$, $x \mapsto (x, F(x))$ is a diffeomorphism with inverse $\pi|_{\Gamma_F}$.

Thus we calculate

$$(\widehat{F})^*((\mathrm{pr}_M)^*\omega_M - (\mathrm{pr}_N)^*\omega_N) = (\mathrm{pr}_M \circ \widehat{F})^*\omega_M - (\mathrm{pr}_N \circ \widehat{F})^*\omega_N = \omega_M - F^*\omega_N,$$

and thus Γ_F is Lagrangian if and only if $\omega_M = F^*\omega_N$ on M, i.e. if and only if F is symplectic.

$\square$

In order to prove that all symplectic manifolds of a fixed dimension $2n$ are locally diffeomorphic to the "symplectic model space" $(\mathbb{R}^{2n}, \omega_0 = \sum_{j=1}^{2n} dx_j \wedge dx_{n+j})$ we introduce the useful tool of "time–dependent vector fields" and the crucial computational formula which is at the base of "Moser's method".

Lemma. *Let M be a manifold and X a vector field on M. Then Cartan's homotopy formula*

$$\mathcal{L}_X \eta = d(i_X \eta) + i_X d\eta \quad \textit{for all} \quad \eta \quad \textit{in} \quad \mathcal{E}^*(M)$$

is equivalent to the formula

$$\frac{d}{dt}((\varphi_t^X)^* \eta) = (\varphi_t^X)^*(d(i_X \eta) + i_X d\eta) \quad \textit{for all} \quad \eta \quad \textit{in} \quad \mathcal{E}^*(M).$$

(Here φ^X denotes of course the flow of X.)

Proof. Evaluating the second formula in $t = 0$ immediately gives Cartan's formula.

By the flow equation $\varphi_{s+t}^X = \varphi_s^X \circ \varphi_t^X$ we find

$$\frac{d}{dt}((\varphi_t^X)^* \eta) = \frac{d}{dt}|_t((\varphi_t^X)^* \eta) = \frac{d}{ds}|_{s=0}((\varphi_{s+t}^X)^* \eta) = (\varphi_t^X)^*(\mathcal{L}_X \eta).$$

Inserting the right hand side of Cartan's formula for $\mathcal{L}_X \eta$ yields the second formula, i.e. the two formulas are equivalent. $\square$

Definition. Let I be a connected interval in $\mathbb{R}$ that contains 0, M a manifold and $\eta : I \times M \to \Lambda^k T^* M$ a smooth map. We define $\eta_t(p) := \eta(t, p)$ for all (t, p) in $I \times M$ and we call η a "time–dependent k–form on M" if η_t is a k–form on M for each t in I.

Lemma. *Let M be a manifold, X a vector field with flow φ^X on M and η a time–dependent k–form on M. Then*

$$\frac{d}{dt}((\varphi_t^X)^* \eta_t) = (\varphi_t^X)^* \Big(d(i_X \eta) + i_X d\eta + \frac{\partial \eta_t}{\partial t} \Big)$$

in all points p in M, t in I, where φ^X is defined. (Here $\frac{\partial \eta_t}{\partial t}$ is of course again a time–dependent k–form on M and thus for fixed t a k–form.)

Proof. The formula follows immediately from the preceding lemma and the Leibniz rule in one variable. $\square$

Definition. Let I be a connected interval in $\mathbb{R}$ and $X : I \times M \to TM$ a smooth map, and let $X_t(p) := X(t, p)$ for all (t, p) in $I \times M$.

(1) We call X a "time–dependent vector field (on M)" if X_t is a vector field on M for each t in M.

(2) Let X be a time–dependent vector field on M. A smooth curve $\gamma : J \to M$ with J open and connected in I is called an "integral curve of the time–dependent vector field X (with initial condition $\gamma(t_0) = p$)" if p is in M, t_0 in J and

$$\dot{\gamma}(t) = (\gamma_*)_t \Big(\frac{d}{dt}\Big|_t \Big) = X_t(\gamma(t)) \quad \text{for all} \quad t \quad \text{in} \quad J \quad \text{and} \quad \gamma(t_0) = p.$$

(3) Let X be a time–dependent vector field on M, $\widehat{M} := I \times M$ and $\widehat{X}(t,p) := \frac{\partial}{\partial t}|_{(t,p)} + X_t(p)$ in $T_{(t,p)}\widehat{M} \cong T_tI \oplus T_pM$. The vector field $\widehat{X}$ on $\widehat{M}$ is called the "suspension of (the time–dependent vector field) X".

Remarks. (1) Typically the interval I is either $\mathbb{R}$ or $[0,1]$, or the latter "with periodic boundary conditions", i.e. $X : S^1 \times M \to TM$ such that X_t is a vector field on M.

(2) Using the suspension $\widehat{X}$ of a time–dependent vector field X it is not difficult to deduce existence and uniqueness of integral curves of X and of maximally defined flow maps $\varphi^X_{t,s}$ such that $\varphi^X_{s,s}(p) = p$ and $\frac{d}{dt}\varphi^X_{t,s}(p) = X_t(\varphi^X_{t,s}(p))$ (where they are defined). The local flow equations are then replaced by $\varphi^X_{s,s} = \mathrm{Id}_M$ and $\varphi^X_{r,t} \circ \varphi^X_{t,s} = \varphi^X_{r,s}$ and in the case that X is time–independent one has $\varphi^X_{t,s} = \varphi^X_{t-s}$. (See, e.g., [AMR] or [Ar2] for more details on time–dependent vector fields and their flows.)

For our purposes the following will be enough:

Proposition. *Let $X : \mathbb{R} \times M \to M$ be a time–dependent vector field such that the closure of $\{p \in M \mid \exists\ t \in \mathbb{R}$ such that $X_t(p) \neq 0\}$ is compact in M. Then the flow maps $\varphi^X_{t,s}$ are defined on all of M for all t, s in $\mathbb{R}$. In particular $\Phi^X_t := \varphi^X_{t,0}$ is a smooth family (in the parameter t in $\mathbb{R}$) of diffeomorphisms of M with $\Phi^X_0 = \varphi^X_{0,0} = Id_M$.*

If furthermore η_t $(t \in \mathbb{R})$ is a smooth time–dependent differential k–form on M then

$$\frac{d}{dt}((\Phi^X_t)^*\eta_t) = (\Phi^X_t)^*\Big(d(i_{X_t}\eta_t) + i_{X_t}d\eta_t + \frac{\partial \eta_t}{\partial t}\Big) \quad \textit{for all} \quad t \quad \textit{in} \quad \mathbb{R}.$$

Proof. Exercise (possibly supported by a textbook as [AMR], [Be] or [GS]). □

Theorem ("Local normal form of symplectic forms on a manifold" or "Theorem of Darboux–Moser–Weinstein"). *Let (M,ω) be a symplectic manifold of dimension $2n$. Then for each point p in M there is an open neighborhood $U = U(p)$ in M and a diffeomorphism $\psi : U \to \psi(U) = V$, V an open set in $\mathbb{R}^{2n}$, such that $\psi^*(\sum_{j=1}^n dx_j \wedge dx_{n+j}) = \omega|_U$.*

Proof. Let $\psi_1 : U_1 \to \psi_1(U_1) = V_1 \subset \mathbb{R}^{2n}$ be any chart such that $\psi_1(p) = 0$. We may assume without loss of generality that V_1 is $\mathbb{R}^{2n}$. Let $\omega_0 = (\psi_1^{-1})^*(\omega|_{U_1})$ and ω_{jk} in $\mathbb{R}$ be defined by $\omega_0(0) = \sum_{j<k}\omega_{jk}(dx_j \wedge dx_k)|_0$. Then ω_0 and $\omega_1 := \sum_{j<k}\omega_{jk}dx_j \wedge dx_k$ are symplectic forms on $\mathbb{R}^{2n}$ fulfilling $\omega_0(0) = \omega_1(0)$.

Since $\mathbb{R}^{2n}$ is contractible there is a 1–form σ on $\mathbb{R}^{2n}$ such that $d\sigma = \omega_0 - \omega_1$. Replacing σ if necessary by $\sigma + df$ with an appropriate function f on $\mathbb{R}^{2n}$ we may furthermore assume that $\sigma(0) = 0$.

Let us define ω_t by $(1-t)\omega_0 + t\omega_1$. Clearly $\omega_t(0) = \omega_0(0)$ and $d\omega_t = 0$. It follows that for $\epsilon_0 > 0$ there is an open neighborhood V_2 of 0 in $\mathbb{R}^{2n}$ such that ω_t is a symplectic form on V_2 for all t in $[-\epsilon_0, 1+\epsilon_0]$. Furthermore there is a $\delta_0 > 0$ such that $\overline{\mathbb{B}_{2\delta_0}(0)} \subset V_2$.

Denoting the inverse of the ismorphism $\omega^\flat : V \to V^*$ on a symplectic vector space (V,ω) by $\omega^\sharp$ we define a smooth time dependent vector field X for x in V_2 (and t in $[-\epsilon_0, 1+\epsilon_0]$) as follows

$$X_t(x) := (\omega_t(x))^\sharp(\sigma(x)).$$

Using a smooth non–negative function χ on $\mathbb{R} \times \mathbb{R}^{2n}$ such that $\chi \equiv 1$ on $[0,1] \times \overline{\mathbb{B}_{\delta_0}(0)}$ and $\chi \equiv 0$ on $\mathbb{R} \times (\mathbb{R}^{2n} \setminus \mathbb{B}_{\frac{3\delta_0}{2}}(0))$ we can extend $\chi \cdot X_t$ to a time–dependent vector field Y_t on $\mathbb{R}^{2n}$ such that $Y_t = X_t$ on $\mathbb{B}_{\delta_0}(0)$ for all t in $[0,1]$ and $Y_t = 0$ for $\|x\| > \frac{3\delta_0}{2}$ and for all t in $\mathbb{R}$.

Thus we can apply the preceding proposition and we have a smooth family $\{\Phi_t := \varphi^Y_{t,0} \,|\, t \in \mathbb{R}\}$ of diffeomorphisms of $\mathbb{R}^{2n}$ with $\Phi_0 = \mathrm{Id}_{\mathbb{R}^{2n}}$, $\Phi_t(0) = 0$ for all t in $\mathbb{R}$ and $\Phi_t(\overline{\mathbb{B}_{2\delta_0}(0)}) = \overline{\mathbb{B}_{2\delta_0}(0)}$ for all t in $\mathbb{R}$. Thus for t in $[0,1]$ the map $\Phi_t = \Phi^X_t$ is a diffeomorphism from $\mathbb{B}_{\delta_0}(0)$ to $\Phi^X_t(\mathbb{B}_{\delta_0}(0))$ such that

$$\frac{d}{dt}(\Phi_t^*\omega_t) = \frac{d}{dt}((\Phi^X_t)^*\omega_t) = (\Phi^X_t)^*(d(i_{X_t}\omega_t) + i_{X_t}d\omega_t + \frac{\partial \omega_t}{\partial t}) = (\Phi^X_t)^*(d\sigma + \omega_1 - \omega_0) = 0.$$

It follows for $t = 1$ that $\Phi := \Phi^X_1 : \mathbb{B}_{\delta_0}(0) \to \Phi^X_1(\mathbb{B}_{\delta_0}(0))$ fulfills $\Phi^*\omega_1 = \omega_0$.

Let furthermore g in $GL(2n, \mathbb{R})$ be such that $(T_g)^*(\sum_{j=1}^n dx_j \wedge dx_{n+j}) = \omega_1$.

We set $\psi := T_g \circ \Phi \circ \psi_1 : U \to V = \psi(U) \subset \mathbb{R}^{2n}$, where $U := \psi_1^{-1}(\mathbb{B}_{\delta_0}(0)) \subset U_1$ is an open neighborhood of p in M. Then ψ is a chart fulfilling $\psi(p) = 0$ and

$$\psi^*\Big(\sum_{j=1}^n dx_j \wedge dx_{n+j}\Big) = \psi_1^* \circ \Phi^* \circ T_g^*\Big(\sum_{j=1}^n dx_j \wedge dx_{n+j}\Big) = \psi_1^*(\omega_0) = \omega|_U.$$

$\square$

Remarks. (1) The above proof relying on the construction of the time–dependent vector field X_t and the formula for $\frac{d}{dt}((\Phi^X_t)^*\omega_t)$ goes back to [M] and is therefore also referred to as "Moser's method". Though the local normal form of symplectic forms on a manifold can be reached in a simpler way we chose this approach since it easily yields proofs for several substantial generalizations. (See e.g. [GS] and [Wei1].)

(2) Local coordinates $(x_1, \ldots, x_{2n})$ as in the preceding theorem, i.e. such that ω is given as $\sum_{j=1}^n dx_j \wedge dx_{n+j}$ are often called "symplectic coordinates". Writing $q_j = x_j, p_j = x_{n+j}$ for $j = 1, \ldots, n$ the form ω is given as $\sum_{j=1}^n dq_j \wedge dp_j$ and we will call such coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ symplectic as well. (The older term for the latter version is "canonical coordinates".)

Bibliographical remarks. *The references cited in the text of this section plus those mentioned at the end of 3.1.*

3.3 Kählerian and almost Kählerian manifolds

Definitions. Let M be a manifold.

(1) A smooth section g of the vector bundle $\otimes^2 T^*M$ over M is called a "pseudo–Riemannian metric (on M)" if the following two conditions are fulfilled for each p in M:

(i) g_p is symmetric, i.e. $g_p(v,w) = g_p(w,v)\ \forall v,w \in T_pM$.

(ii) g_p is non–degenerate, i.e. for all $0 \neq v$ in T_pM there is a w in T_pM such that $g_p(v,w) \neq 0$.

(2) A pseudo–Riemannian metric on M is called "Riemannian" if $g(v,v) > 0\ \forall v \in T_pM \setminus \{0\}$ for all p in M.

(3) A pair (M,g) consisting of a manifold and a (pseudo-)Riemannian metric is called a "(pseudo-)Riemannian manifold".

Remark. A (pseudo-)Riemannian metric on a manifold is nothing else than a smoothly varying assignement of a (pseudo-)Riemannian metric to each tangent space.

Examples.

(1) Let M be $\mathbb{R}^m$ and $g = \sum_{1\le i,j\le m} g_{ij} e_i^* \otimes e_j^*$ be a (pseudo-)Riemannian metric on the vector space $V = \mathbb{R}^m$. Setting for each p in M $g_p = \sum_{1\le i,j\le m} g_{ij}(dx_i)_p \otimes (dx_j)_p$ (where $(x_1,\ldots,x_m)$ are the canonical global coordinates on M), we get a (pseudo-)Riemannian metric on M. Though slightly abusive, it is convenient to denote this (pseudo-)Riemannian metric by the letter g as well.

(2) Let (M,g) be a Riemannian manifold and $N \subset M$ a closed submanifold. Then for each p in N T_pN is a subspace of T_pM and thus $g_p|_{T_pN}$ is a Riemannian metric on the vector space T_pN. It is easy to check that $(N, g|_{TN})$ is a Riemannian manifold.

Definition. Let M be a (real!) manifold and $J : TM \to TM$ be a smooth vector bundle homomorphism over Id_M such that $J^2 = J \circ J = -\mathrm{Id}_{TM}$. Then J is called an "almost–complex structure on M" and the pair (M,J) is called an "almost–complex manifold".

Proposition. *Let M be a real manifold and $\mathfrak{A}$ a holomorphic atlas on M. Then M carries an almost–complex structure canonically associated to $\mathfrak{A}$. Furthermore, if $\varphi : U \to V = \varphi(U) \subset \mathbb{C}^m$ is a chart on an open subset U of M that is (holomorphically!) compatible with $\mathfrak{A}$ and if $\varphi = {}^t(z_1,\ldots,z_m)$ with $z_k = x_k + \sqrt{-1}y_k$ for $k = 1,\ldots,m$, then*

$$J\Big(\frac{\partial}{\partial x_k}\Big) = \frac{\partial}{\partial y_k} \quad \textit{and} \quad J\Big(\frac{\partial}{\partial y_k}\Big) = -\frac{\partial}{\partial x_k} \quad \textit{for } k = 1,\ldots,m$$

in these coordinates.

Proof. The complex–analytic atlas $\mathfrak{A}$ provides the vector bundle $TM \to M$ with the structure of a holomorphic vector bundle and in particular each fiber is a complex vector space. For each p in M, one defines J_p as the real–linear endomorphism induced on the real tangent space T_pM of the underlying real manifold M by the multiplication with the complex number $i = \sqrt{-1}$ on the space T_pM viewed as a complex vector space. It follows that J is an almost–complex structure on the real manifold M, canonically associated to the complex–analytic atlas $\mathfrak{A}$.

Without loss of generality we may now assume that $M = V$ is open in $\mathbb{C}^m$ and $\varphi = \mathrm{Id}_V = {}^t(z_1,\ldots,z_m)$ with $z_k = x_k + iy_k$ and x_k, y_k real. Considering $\frac{\partial}{\partial x_k}$ and $\frac{\partial}{\partial y_k}$ as elements of the complex vector space T_pM (for p any point in M), we find

$$dz_j(\frac{\partial}{\partial x_k}) = (dx_j + i\,dy_j)(\frac{\partial}{\partial x_k}) = \delta_{j,k} \quad \text{and} \quad dz_j(\frac{\partial}{\partial y_k}) = i\,\delta_{j,k}.$$

Using the $\mathbb{C}$–linearity of the functionals dz_j on T_pM we find

$$dz_j(i\cdot\frac{\partial}{\partial x_k}) = (i\,dz_j)(\frac{\partial}{\partial x_k}) = i\,\delta_{j,k}.$$

It follows that, for $k = 1, \ldots, m$,

$$J\Big(\frac{\partial}{\partial x_k}\Big) = \frac{\partial}{\partial y_k} \quad \text{and} \quad J\Big(\frac{\partial}{\partial y_k}\Big) = -\frac{\partial}{\partial x_k}.$$

□

Definition. An almost–complex structure J on a real manifold M that is canonically associated to a complex–analytic atlas $\mathfrak{A}$ (as in the preceding proposition) is called a "complex structure".

Remarks.

(1) Though the distinction between complex structures and almost–complex structures is important, some texts on symplectic geometry are not attentive to it.

(2) An almost–complex structure that fulfills a certain "integrability" condition is called an "integrable almost–complex structure". In finite dimensions every integrable smooth almost–complex structure is already a complex structure by a deep theorem of Newlander and Nirenberg (see the original work [NN] or [H] for a proof).

Definitions. Let (M, J) be an almost–complex manifold and g a (pseudo-)Riemannian metric on M.

(1) The metric g is called "almost (pseudo-)Hermitian" if, for all p in M

$$g_p(J_pv, J_pw) = g_p(v, w)\,\text{for all}\,v, w\,\text{in}\,T_pM.$$

(2) If J is a complex structure and g almost (pseudo-)Hermitian then g is called "(pseudo-) Hermitian".

(3) If g is almost (pseudo-)Hermitian, the 2–form ω, defined by

$$\omega_p(v, w) = g_p(J_p(v), w) \quad \forall\ p \in M, \quad \forall v, w \in T_pM,$$

is called the "fundamental 2–form (on the almost (pseudo-)Hermitian manifold (M, J, g))".

Lemma. *Let (M, J, g) be an almost pseudo–Hermitian manifold. Then the fundamental 2–form is almost–symplectic.*

Proof. We only need to check that ω_p is alternating and non–degenerate for all p in M. This is proven in Section 1.5 since (T_pM, J_p, g_p) is a pseudo–Hermitian vector space for all p in M. □

Definitions. Let (M, J, g) be an almost (pseudo-)Hermitian manifold and ω its fundamental 2–form.

(1) The triple (M, J, g) is called an "almost (pseudo-)Kählerian manifold" if ω is closed.

(2) An almost (pseudo-)Kählerian manifold is called "(pseudo-)Kählerian" if J is a complex structure.

Remarks.

(1) The last lemma implies that each almost pseudo–Kählerian manifold is symplectic.

(2) Obviously every Kählerian or pseudo–Kählerian manifold is almost pseudo–Kählerian and thus, a fortiori, symplectic.

(3) It was conjectured that each symplectic manifold is Kählerian. This is wrong (see example (5) below) but a "partial converse" of (2) holds true:

Proposition. *Let (M, ω) be a symplectic manifold and $\mathcal{J}_\omega(M)$ be the set of almost–complex structures J on M such that (M, J, g_J) is almost Kählerian with fundamental 2–form equal to ω. Then $\mathcal{J}_\omega(M)$ is not empty, i.e. every symplectic manifold is almost Kählerian.*

Remark. Upon considering $\mathcal{J}_\omega(M)$ as a subset of the "Fréchet space" $\Gamma_{C^\infty}(M, \mathrm{End}(TM))$ one can use Section 1.5 to prove the following important sharpening of the preceding proposition: the topological space $\mathcal{J}_\omega(M)$ is non–empty and continuously contractible to a point. (See, e.g., [McDS].)

Proof of the proposition. The existence of a partition of unity subordinate to an appropriate covering of M easily shows that M carries a Riemannian metric g.

Applying the theorem of Section 1.5 to $V = T_pM$ yields maps

$$\Psi_p : \mathcal{R}(T_pM) \to \mathcal{J}_{\omega_p}(T_pM)$$

such that $\Psi_p(g_p) = J_p$ and (T_pM, J_p, g_p) is Hermitian for all p in M. Since Ψ is real–analytic in the variable $g \in \mathcal{R}(V)$ one easily poves, by going to local charts, that J is a smooth section of End(TM). Thus J is an almost–complex structure on M. The theorem in Section 1.5 implies furthermore that the Riemannian metric g_J defined by

$$(g_J)_p(v, w) = \omega_p(v, J_p(w)) \quad \forall\, p \in M, \quad \forall v, w \in T_pM$$

is almost Hermitian on (M, J) and the fundamental 2–form of (M, J, g_J) is equal to ω. $\square$

Examples.

(1) Let $(M, \mathfrak{A})$ be a 2–dimensional real manifold with a complex–analytic atlas and g a Hermitian metric on M. Then (M, J, g) is Kählerian since every 2–form on M is closed.

(2) Let $M = \mathbb{C}^n \cong \mathbb{R}^{2n}$ and $g = \sum_{k=1}^n (dx_k \otimes dx_k + dy_k \otimes dy_k)$ the standard Riemannian metric on $\mathbb{R}^{2n}$. Since $J(\frac{\partial}{\partial x_k}) = -\frac{\partial}{\partial y_k}$ the metric g is Hermitian and the fundamental 2–form of (M, J, g) is given as follows

$$\omega = \sum_{k=1}^n dx_k \wedge dy_k.$$

Obviously, ω is closed and thus (M, J, g) is Kählerian.

(3) Given a discrete subgroup Γ of $(\mathbb{C}^n, +)$ one easily checks that the complex structure, the metric and thus the fundamental 2–forms "descend" to the quotient $\mathbb{C}^n/\Gamma = \Gamma \backslash \mathbb{C}^n$,

since they are in fact invariant under the whole group $(\mathbb{C}^n, +)$ acting on $M = \mathbb{C}^n$. Thus $\mathbb{C}^n/\Gamma$ is Kählerian.

(4) Open subsets of Kählerian manifolds, products of Kählerian manifolds and complex submanifolds of Kählerian manifolds are naturally equipped with induced complex structures and Hermitian metrics such that they are Kählerian.

(5) Let $\widetilde{M} = N_{\mathbb{R}}$ be as in Example (3.2) of Section (2.2), i.e.,

$$\widetilde{M} = \left\{ \left(\begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix}, w \right) \middle| \, x, y, z, w \in \mathbb{R} \right\} \cong \mathbb{R}^3 \times \mathbb{R}.$$

Let $\widetilde{\omega} = dy \wedge dz + dx \wedge dw$, then $\widetilde{\omega}$ is a symplectic form on $\widetilde{M}$ that is invariant under the action of $N_{\mathbb{R}}$ on $\widetilde{M}$ given by left–multiplication. It follows that $M = N_{\mathbb{Z}} \backslash N_{\mathbb{R}}$ is a compact manifold with a unique 2–form such that under $\pi : \widetilde{M} \to M$ we have $\pi^*(\omega) = \widetilde{\omega}$. It follows easily that (M, ω) is symplectic since π has everywhere rank four.

Standard arguments from algebraic topology show that the dimension of $H^1_{dR}(M, \mathbb{R})$ is equal to three which shows that M cannot carry a Kählerian metric. (See [BT] and [GriHa] for more details on the algebraic topology respectively Hodge theory needed to show the above assertions.) The manifold M was considered by Thurston to exhibit a compact symplectic manifold allowing no Kählerian metric. (See, e.g., [Wei2].)

(6) The complex projective space $\mathbb{P}_n(\mathbb{C})$ is a very important Kählerian manifold with respect to the so-called "Fubini-Study metric" and its associated fundamental 2-form, the "Fubini-Study form" ω_{FS}. (See, e.g., [GriHa] or [J] for more details.)

Bibliographical remarks. *Beside the references cited in the text we would like to mention [MK] and [Wel] for the theory of Kählerian manifolds. The reader should be aware that traditionally the class of Kählerian manifolds is viewed as a special case of complex manifolds and not of symplectic manifolds and thus notations are rather "complex" than "real".*

3.4 Hamiltonian dynamical systems on symplectic manifolds

Remark. Given an almost–symplectic manifold (M, ω) the map

$$\omega^\flat : TM \to T^*M, \omega^\flat(v_p) := \omega^\flat_p(v_p) \quad \text{for all } p \text{ in } M \text{ and all } v_p \text{ in } T_pM$$

is a vector bundle isomorphism over Id_M, the inverse of which we denote by $\omega^\sharp$.

Definitions. (1) Let (M, ω) be a symplectic manifold and H a smooth function on M. Then we define a vector field X_H on M by

$$X_H = \omega^\sharp(dH).$$

The vector field X_H is called the "Hamiltonian vector field associated to the Hamilton function H" (or "symplectic gradient of H").

(2) A triple (M, ω, H) consisting of a symplectic manifold (M, ω) and a smooth function H on M is called a "Hamiltonian dynamical system".

Remark. Since we associated to a function H on (M, ω) a vector field X_H, a Hamiltonian dynamical system comes equipped with the (local) flow φ^{X_H}, i.e. a local $\mathbb{R}$–action on M. This explains the terminology.

Remark. Since $\omega^\flat \circ \omega^\sharp = \mathrm{Id}_{T^*M}$ the vector field X_H is often defined as the unique vector field on M fulfilling

$$\omega(X_H, \cdot) = dH.$$

This formula is clearly equivalent to the above definition and is in fact very useful in computations.

Proposition. *Let (M, ω, H) be a Hamiltonian dynamical system and $\varphi = (q_1, \ldots, q_n, p_1, \ldots, p_n) : U \to \varphi(U) = V \subset \mathbb{R}^{2n}$ a symplectic chart defined on an open subset U of M. Then the Hamiltonian vector field is given (on V) by the following formula*

$$X_H = \sum_{j=1}^{n} \left(\frac{\partial H}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial H}{\partial q_j} \frac{\partial}{\partial p_j} \right).$$

Remarks.

(1) It is understood that the function H should be read as $\widetilde{H} = H \circ \varphi^{-1}$ in the formula of the proposition. We follow the usual practice to suppress this inconvenient notation.

(2) From the last proposition in Section 2.5 we know that (on V) the differential equation ("Hamilton's equations") for the flow of X_H as in the formula of the proposition is then given by

$$\dot{q}_j = \frac{\partial H}{\partial p_j}, \ \dot{p}_j = -\frac{\partial H}{\partial p_j} \quad \text{for} \quad j = 1, \ldots, n.$$

Proof of the proposition. Let first $\varphi : U \to V \subset \mathbb{R}^{2n}$ be any chart on M and $\widetilde{H} = H \circ \varphi^{-1}, \widetilde{\omega} = (\varphi^{-1})^* \omega$. Then $\widetilde{\omega}$ is a symplectic form on V and the vector field $X_{\widetilde{H}} = \widetilde{\omega}^\sharp(d\widetilde{H})$ fulfills

$$\omega(X_H, v) = \widetilde{\omega}(X_{\widetilde{H}}, \varphi_* v) \quad \forall\, v \in T_m M \quad \forall\, m \in U.$$

It follows by the non–degeneracy of ω and $\tilde{\omega}$ that $(\varphi_*)_m X_H(m) = X_{\widetilde{H}}(\varphi(m))$ for all m in U, i.e., X_H is given by $X_{\widetilde{H}}$ in the chart φ.

Let us now assume that $\varphi = (q_1, \ldots, q_n, p_1, \ldots, p_n)$ is a symplectic chart, i.e. $\widetilde{\omega} = (\varphi^{-1})^* \omega = \sum_{j=1}^{n} dq_j \wedge dp_j =: \omega_0$ on $V \subset \mathbb{R}^{2n}$. Let $j \in \{1, \ldots, n\}$ and $X_{\widetilde{H}} = \sum_{k=1}^{n} \left(\alpha_k \frac{\partial}{\partial q_k} + \beta_k \frac{\partial}{\partial p_k} \right)$ with α_k, β_k in $\mathcal{E}(V)$. Then

$$\alpha_j = \omega_0 \Big(X_{\widetilde{H}}, \frac{\partial}{\partial p_j} \Big) = d\widetilde{H} \Big(\frac{\partial}{\partial p_j} \Big) = \frac{\partial \widetilde{H}}{\partial p_j} \text{ and } \beta_j = -\omega_0 \Big(X_{\widetilde{H}}, \frac{\partial}{\partial q_j} \Big) = -d\widetilde{H} \Big(\frac{\partial}{\partial q_j} \Big) = -\frac{\partial \widetilde{H}}{\partial q_j}.$$

It follows that

$$X_{\widetilde{H}} = \sum_{j=1}^{n} \Big(\frac{\partial \widetilde{H}}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial \widetilde{H}}{\partial q_j} \frac{\partial}{\partial p_j} \Big),$$

as claimed in the proposition. □

Let us observe that there is another way of expressing Hamilton's equations on $\mathbb{R}^{2n}$ (in fact, more generally, on almost Kählerian manifolds) which brings into play the almost–complex structure:

Lemma. *Let V be open in $\mathbb{R}^{2n}$, H a smooth function on V and $\mathbb{J}_n = \begin{pmatrix} 0 & -\mathbb{E}_n \\ \mathbb{E}_n & 0 \end{pmatrix}$ the standard almost–complex structure on $\mathbb{R}^{2n}$ (as in Section 1.5). Then Hamilton's equations are equivalent to*

$$\dot{x} = -\mathbb{J}(\nabla H(x)).$$

Proof. Let $x = (q_1, \ldots, q_n, p_1, \ldots, p_n) =: (q, p)$.

Then Hamilton's equations

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix}$$

are obviously equivalent to

$$\dot{x} = \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H}{\partial q} \\ \frac{\partial H}{\partial p} \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{pmatrix}.$$

□

Definition. Let (M, ω) a symplectic manifold and X in $\mathfrak{X}(M)$.

(1) We call X a "symplectic vector field" if $\mathcal{L}_X \omega = 0$.

(2) We call X a "Hamiltonian vector field" if there exists a smooth function H on M such that $X = X_H$.

Lemma. *Let (M, ω) be a symplectic manifold and X in $\mathfrak{X}(M)$. Then $\mathcal{L}_X \omega = 0$ if and only if*

$$(\varphi_t^X)^* \omega = \omega,$$

in all points where the last equation makes sense.

Proof. Obviously the second formula implies the first. The equality

$$\frac{d}{dt}((\varphi_t^X)^* \omega) = (\varphi_t^X)^* (\mathcal{L}_X \omega)$$

which follows from the flow equations (compare Section 3.2) shows that $\mathcal{L}_X \omega = 0$ implies $(\varphi_t^X)^* \omega = \omega$. Thus the two formulas are equivalent. □

Lemma. *Let (M, ω) be a symplectic manifold and X a symplectic vector field. Then the powers of ω are invariant under X, i.e.,*

$$\mathcal{L}_X(\omega^k) = 0 \quad \textit{for} \quad k = 0, 1, \ldots, (\dim_{\mathbb{R}} M)/2.$$

Proof. Obvious. $\square$

Corollary. *Let (M, ω) be a symplectic manifold and X a symplectic manifold. Let furthermore t be in $\mathbb{R}$ and U open in M such that φ_t^X is defined on U. Then for all $(2k)$–dimensional orientable submanifolds of U one has that*

$$\int_{\Sigma_{2k}} \omega^k = \pm \int_{\varphi_t^X(\Sigma_{2k})} \omega^k$$

(if one and then both sides of the equality is finite).

Proof. Let us without loss of generality assume that $\varphi_t^X : \Sigma_{2k} \to \varphi_t^X(\Sigma_{2k})$ is orientation preserving. It follows that $\int_{\varphi_t^X(\Sigma_{2k})} \omega^k = \int_{\Sigma_{2k}} (\varphi_t^X)^*(\omega^k) = \int_{\Sigma_{2k}} \omega^k$ by the invariance of the integral under orientation preserving diffeomorphisms. $\square$

Remark. Assuming that $\dim_{\mathbb{R}} M = 2n$ and taking $\Sigma_{2k} = U$ an open set with finite "phase volume" $\int_U \Omega = \int_U \left(\frac{(-1)^{\frac{n(n-1)}{2}}}{n!} \right) \omega^n$ yields the result that "the phase volume is invariant under symplectic flows".

Lemma. *Let (M, ω) be a symplectic manifold and X in $\mathfrak{X}(M)$. Then*

(i) X is symplectic if and only if the one–form $i_X\omega$ is closed, and

(ii) X is Hamiltonian if and only if $i_X\omega$ is exact.

In particular, a Hamiltonian vector field is symplectic.

Proof. Since ω is closed, Cartan's homotopy formula implies for all X in $\mathfrak{X}(M)$ that $\mathcal{L}_X\omega = d(i_X\omega)$. The assertions follow now immediately. $\square$

Since a closed form on a manifold is always locally exact by Poincaré's lemma the following notions are rather natural:

Definition. Let (M, ω) be a symplectic manifold.

(1) A symplectic vector field is also called a "locally Hamiltonian vector field."

(2) The set of all Hamiltonian vector fields (respectively all locally Hamiltonian vector fields) is denoted by $\mathrm{Ham}(M, \omega)$ (respectively $\mathrm{Ham}_{\mathrm{loc}}(M, \omega)$).

Lemma. *Let (M, ω) be a symplectic manifold and let X and Y be locally Hamiltonian vector fields. Then $[X, Y]$ is the Hamiltonian vector field associated to the smooth function $H = -\omega(X, Y)$.*

Proof. It is enough to show that

$$dH(Z) = \omega([X, Y], Z) \quad \text{for all } Z \quad \text{in } \mathfrak{X}(M).$$

Using the formula $i_{[A,B]} = [\mathcal{L}_A, i_B]$ on $\mathcal{E}^*(M)$ for A, B in $\mathfrak{X}(M)$ we find:

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

Closedness of ω implies

$$0 = X(\omega(Y,Z)) - Y(\omega(X,Z)) + Z(\omega(X,Y)) - \omega([X,Y],Z) + \omega([X,Z],Y) - \omega([Y,Z],X).$$

Combining these identities yields

$$2\omega([X,Y],Z) = -Z(\omega(X,Y)) + \omega([X,Y],Z), \text{ i.e., } \quad \omega([X,Y],Z) = -Z(\omega(X,Y)) = dH(Z).$$

□

Corollary. *Let (M,ω) be a symplectic manifold. Then $Ham_{loc}(M,\omega)$ is a Lie subalgebra of $(\mathfrak{X}(M),[,])$ fulfilling $[Ham_{loc}(M,\omega), Ham_{loc}(M,\omega)] \subset Ham(M,\omega)$. Furthermore $Ham(M,\omega)$ is a Lie subalgebra of $(\mathfrak{X}(M),[,])$ and an ideal in $Ham_{loc}(M,\omega)$.*

Remark. A subspace $\mathfrak{h}$ of a Lie algebra $(\mathfrak{g},[,])$ is called an "ideal" if $[X,H]$ is in $\mathfrak{h}$ for all X in $\mathfrak{g}$ and H in $\mathfrak{h}$.

Proof of the corollary. Obvious from the preceding lemma. □

Exercise. Let (M,ω) be a symplectic manifold. Then the quotient vector space $\mathrm{Ham}_{\mathrm{loc}}(M,\omega)/\mathrm{Ham}(M,\omega)$ is canonically isomorphic to $H^1_{dR}(M)$. Supplying the latter with the trivial commutator, i.e. all brackets are zero, the following sequence of Lie algebra morphisms is exact:

$$\{0\} \to \mathrm{Ham}(M,\omega) \xrightarrow{\alpha} \mathrm{Ham}_{\mathrm{loc}}(M,\omega) \xrightarrow{\beta} H^1_{dR}(M) \to \{0\},$$

where α is the natural injection and β the projection on the above mentioned quotient followed by the canonical isomorphism.

Definition. Let (M,ω) be an almost–symplectic manifold and let H, H_1, H_2 be in $\mathcal{E}^0(M)$.

(1) The "almost–symplectic gradient of H" is the vector field $X_H = \omega^\sharp(dH)$.

(2) The "Poisson bracket of H_1 and H_2" is the smooth function $\omega(X_{H_1}, X_{H_2})$ denoted by $\{H_1, H_2\}$.

Lemma. *Let (M,ω) be an almost–symplectic manifold and $H_1, H_2 \in \mathcal{E}^0(M)$. Then*

(i) $\{H_1, H_2\} = -X_{H_1}(H_2) = X_{H_2}(H_1)$, *and*

(ii) the map

$$\{,\} : \mathcal{E}^0(M) \times \mathcal{E}^0(M) \to \mathcal{E}^0(M), (H_1, H_2) \mapsto \{H_1, H_2\}$$

is $\mathbb{R}$–bilinear and anti–symmetric, and fulfills

$$\{H_1, H_2 \cdot H_3\} = \{H_1, H_2\} \cdot H_3 + H_2 \cdot \{H_1, H_3\}.$$

Proof. We have

$$\{H_1, H_2\} = \omega(X_{H_1}, X_{H_2}) = i_{X_{H_2}}(\omega(X_{H_1}, \cdot)) = i_{X_{H_2}}(dH_1) = X_{H_2}(H_1).$$

Since ω is anti–symmetric the first assertion is thus proven.

Bilinearity over $\mathbb{R}$ and anti–symmetry of $\{,\}$ follow directly from the properties of an almost–symplectic form.

It remains to show that for each H_1 in $\mathcal{E}^0(M)$ the map $\{H_1, \}: \mathcal{E}^0(M) \to \mathcal{E}^0(M)$ is a derivation:

$$\begin{aligned}\{H_1, H_2 \cdot H_3\} &= -X_{H_1}(H_2 \cdot H_3) = -X_{H_1}(H_2) \cdot H_3 - H_2 \cdot X_{H_1}(H_3) \\ &= \{H_1, H_2\} \cdot H_3 + H_2 \cdot \{H_1, H_3\}.\end{aligned}$$

□

One important motivation of the closedness condition on ω is the following

Proposition. *Let (M, ω) be an almost–symplectic manifold. Then the Poisson bracket $\{,\}$ fulfills the Jacobi–identity on $\mathcal{E}^0(M)$ if and only if ω is closed.*

Proof. For F, G and H functions on an almost–symplectic manifold one has by a simple calculation

$$\omega([X_F, X_G], X_H) = -\{F, \{G, H\}\} + \{G, \{F, H\}\}.$$

A lengthy but elementary calculation shows now that

$$(d\omega)(X_{H_1}, X_{H_2}, X_{H_3}) = \{H_1, \{H_2, H_3\}\} - \{\{H_1, H_2\}, H_3\} - \{H_2, \{H_1, H_3\}\}.$$

Thus closedness of ω implies that the Jacobi–identity holds for $\{,\}$ on $\mathcal{E}^0(M)$.

On the other hand given an element φ in T_p^*M for a p in M there exists a smooth function H on M such that $\varphi = (dH)(p)$. Thus by the non–degeneracy of an almost–symplectic form there exists, given a point p in M and v_1, v_2, v_3 in T_pM three functions H_1, H_2, H_3 on M such that $X_{H_j}(p) = v_j$ for $j = 1, 2, 3$. Assuming now that $\{,\}$ fulfills the Jacobi–identity we conclude that the three–form $(d\omega)$ satisfies the following condition:

$$(d\omega)_p(v_1, v_2, v_3) = 0 \quad \text{for all } p \quad \text{in } M \quad \text{and for all } v_1, v_2, v_3 \quad \text{in } T_pM,$$

i.e., $d\omega = 0$. □

Definition. Let (M, ω) be a symplectic manifold and $\kappa : \mathcal{E}(M) \to \mathrm{Ham}(M, \omega)$ be defined by $\kappa(H) = -X_H$. The following sequence of $\mathbb{R}$–vector spaces and $\mathbb{R}$–linear maps is called the "fundamental sequence (on a symplectic manifold)":

$$\{0\} \to \ker\kappa \xrightarrow{j} \mathcal{E}(M) \xrightarrow{\kappa} \mathrm{Ham}(M, \omega) \to \{0\}.$$

(The map j is the injection of $\ker\kappa$ in $\mathcal{E}(M)$.)

Lemma. *The fundamental sequence on a symplectic manifold is an exact sequence of Lie algebras and* $\ker\kappa$ *is the space of locally constant functions on M.*

Remark. A continuous function on a manifold is called locally constant if for each point of the manifold there is a neighborhood of this point such that the function is constant on this neighborhood. Obviously a continuous function is locally constant if and only if it is constant on each connected component of the manifold. A C^1–function is thus locally constant if and only if $df = 0$.

Proof of the lemma. Since ω is non–degenerate $\kappa(H) = -X_H$ is zero for H in $\mathcal{E}(M)$ if and only if $dH = 0$, i.e., $\ker\kappa$ is the space of locally constant functions on M and the Poisson bracket of such two functions is the zero function. Thus $\ker\kappa$ is a Lie subalgebra of $(\mathcal{E}(M), \{,\})$. which is in fact easily seen to be an abelian ideal.

Since κ is $\mathbb{R}$–linear and surjective it remains only to show that $\kappa(\{H_1, H_2\}) = [\kappa(H_1), \kappa(H_2)]$ for all H_1, H_2 in $\mathcal{E}(M)$. Since vector fields on a (finite dimensional) manifold can be identified with derivations it is enough to prove that both act in the same way on functions. Let thus H_3 be in $\mathcal{E}(M)$, then

$$\begin{aligned}[\kappa(H_1), \kappa(H_2)](H_3) &= X_{H_1}(X_{H_2}(H_3)) - X_{H_2}(X_{H_1}(H_3))\\ &= -(\{\{H_3, H_1\}, H_2\} + \{H_1, \{H_3, H_2\}\}).\end{aligned}$$

By the Jacobi–identity the last right–hand side equals

$$-\{H_3, \{H_1, H_2\}\} = -X_{\{H_1,H_2\}}(H_3) = \kappa(\{H_1, H_2\})(H_3),$$

showing the assertion. □

Definition. Let (M, ω, H) be a Hamiltonian dynamical system and F a smooth function on M. The function F is called a "first integral (of the motion)" or a "conserved quantity" if and only if F is constant on the integral curves of H.

Lemma. *Let (M, ω, H) be a Hamiltonian dynamical system and F in $\mathcal{E}(M)$. Then F is a first integral if and only if $\{H, F\} = 0$.*

Proof. Considering F as a 0–form on M we have

$$\frac{d}{dt}(F(\varphi_t^{X_H})) = \frac{d}{dt}((\varphi_t^{X_H})^* F) = (\varphi_t^{X_H})^*(\mathcal{L}_{X_H} F) = -((\varphi_t^{X_H})^*\{H, F\}).$$

Thus $\{H, F\} = 0$ if and only if $F \circ \varphi_t^{X_H} = F \circ \varphi_0^{X_H} = F$, i.e., if and only if F is constant on the integral curves of X_H. □

Proposition ("Noether's theorem"). *Let Q be a manifold, $(M, \omega) = (T^*Q, \omega^{T^*Q})$, and H a smooth function on M. Let furthermore $\varphi : \Omega \to Q$ (Ω open in $\mathbb{R} \times Q$) be a local flow on Q such that the induced local flow $\widehat{\varphi}$ on M fulfills $H(\widehat{\varphi}_t(m)) = H(m)$ whenever the left–hand side is defined. Then there exists a smooth function F on M such that*

$$\{H, F\} = 0 \quad \textit{and} \quad \varphi^{X_F} = \widehat{\varphi}.$$

Proof. Let X be the vector field on M that generates the local flow $\widehat{\varphi}$, i.e., $X(m) = \frac{d}{dt}|_0 \widehat{\varphi}_t(m)$ for m in M. We define a smooth function F on M by setting $F = \theta^{T^*Q}(X)$. It follows that ($\theta = \theta^{T^*Q}, \omega = \omega^{T^*Q} = -d\theta$):

$$dF = (d \circ i_X)\theta = -(i_X \circ d)\theta + \mathcal{L}_X \theta = i_X \omega = \omega^\flat(X)$$

by the fact that the "lift" $\widehat{\varphi}$ of the point transformations φ preserves the 1–form θ. Thus $X = X_F$, the Hamiltonian vector field associated to F and $\widehat{\varphi}_t = \varphi_t^{X_F}$. Obiously $H \circ \varphi_t^{X_F} = H \circ \widehat{\varphi}_t = H$ and therefore $\{H, F\} = 0$. □

Remarks. (1) The analogous statement for "Lagrangian dynamical systems" on TQ is in fact more often referred to as Noether's theorem.

(2) In physics the above theorem is often formulated as follows: "a continuous symmetry of the Hamiltonian gives a conserved quantity". (The word "continuous" should be viewed in distinction to "discrete" here, i.e. a "continuous symmetry" is a local $\mathbb{R}$–action that preserves ω and H.)

Bibliographical remarks. *As at the end of Section 3.1.*

References

[AM] R. Abraham and J. Marsden, *Foundations of Mechanics*, Benjamin/Cummings, Reading 1967 & 1978.

[AMR] R. Abraham, J. Marsden and T. Ratiu, *Manifolds, tensor analysis, and applications*, Addison-Wesley, Reading 1983 & Springer, New York-Berlin 1988.

[Ar1] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, New York 1978.

[Ar2] V.I. Arnold, *Ordinary differential equations*, Springer, Berlin 1992.

[Be] R. Berndt, *Einführung in die Symplektische Geometrie*, Vieweg, Braunschweig-Wiesbaden 1998. English translation: *An introduction to symplectic geometry*, AMS, Providence RI 2000.

[Bo] W.M. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, Academic Press, Orlando 1975 & 1986.

[BT] R. Bott and L. Tu, *Differential forms in algebraic topology*, Springer, New York 1982.

[Bre] G. Bredon, *Introduction to compact transformation groups*, Academic Press, New York-London, 1972.

[Brö] T. Bröcker, *Analysis in mehreren Variablen (einschließlich gewöhnlicher Differentialgleichungen und des Satzes von Stokes)*, Teubner, Stuttgart 1980.

[BröJä] T. Bröcker and K. Jänich, *Introduction to differential topology*, Cambridge University Press, Cambridge-New York 1982.

[Bry] R. Bryant, An introduction to Lie groups and symplectic geometry. In: D. Freed et al. (Editors), *Geometry and quantum field theory*, AMS, Providence RI 1995.

[CM] P.R. Chernoff and J.E. Marsden, *Properties of infinite dimensional Hamiltonian systems*, Lecture Notes in Mathematics, Vol. 425. Springer, Berlin-New York, 1974.

[GHL] S. Gallot, D. Hulin and J. Lafontaine, *Riemannian Geometry*, Springer, Heidelberg 1990.

[God] C. Godbillon, *Géométrie Différentielle et Mécanique Analytique*, Hermann, Paris 1969.

[Gol] H. Goldstein, *Classical mechanics*, Addison-Wesley, Reading Mass. 1980.

[Gre] W. Greub, *Multilinear algebra*, Springer, New York 1978.

[GriHa] P. Griffiths and J. Harris, *Principles of algebraic geometry*, John Wiley, New York, 1994.

[GS] V. Guillemin and S. Sternberg, *Symplectic techniques in physics*, Cambridge University Press, Cambridge, 1984.

[H] L. Hörmander, *An introduction to complex analysis in several variables*, North-Holland, Amsterdam-New York, 1990.

[J] J. Jost, *Riemannian geometry and geometric analysis*, Springer, Berlin 1995 & 1998.

[KL] M. Karoubi and C. Leruste, *Algebraic topology via differential geometry*, Cambridge University Press, Cambridge-New York, 1987.

[Lan] S. Lang, *Differential manifolds*, Addison-Wesley, Reading, Mass. 1972.

[Lau] R. Lauterbach, *Klassische Mechanik und Hamiltonische Systeme*, http://www.math.uni-hamburg.de/home/lauterbach/vorlesung.html

[McDS] D. McDuff and D. Salamon, *Introduction to symplectic topology*, Oxford University Press, New York, 1995.

[MK] J. Morrow and K. Kodaira, *Complex manifolds*, Holt, Rinehart and Winston, New York-Montreal 1971.

[M] J. Moser, On the volume elements on a manifold, *Trans. Amer. Math. Soc.* 120, 1965, 286–294.

[NN] A. Newlander and L. Nirenberg, Complex analytic coordinates in almost complex manifolds, *Annals of Mathematics* (2) 65, 1957, 391–404.

[Sch] F. Scheck, *Mechanics. From Newton's laws to deterministic chaos*, Springer, Berlin 1999.

[Sp] M. Spivak, *A comprehensive introduction to differential geometry. Vol. I*, Publish or Perish, Berkeley 1979.

[Wei1] A. Weinstein, Symplectic manifolds and their Lagrangian submanifolds, *Advances in Mathematics*, 6, 1971, 329–346.

[Wei2] A. Weinstein, Lectures on symplectic manifolds, *CBMS Regional Conference Series in Mathematics*, 29, American Mathematical Society, Providence, R.I., 1977 & 1979.

[Wel] R. O. Wells, *Differential analysis on complex manifolds*, Springer, New York-Berlin 1980.

[Wu] T. Wurzbacher, Fermionic second quantization and the geometry of the Hilbert space Grassmannian, *DMV-Seminar on "Infinite-dimensional Kähler manifolds" Oberwolfach November 1995*, Preprint 1998.
http://www-irma.u-strasbg.fr/~wurzbach/publications.html