CHAPTER I. Manifolds and Vector Fields

1. Differentiable Manifolds

1.1. Manifolds. A *topological manifold* is a separable metrizable space M which is locally homeomorphic to $\mathbb{R}^n$. So for any $x \in M$ there is some homeomorphism $u : U \to u(U) \subseteq \mathbb{R}^n$, where U is an open neighborhood of x in M and $u(U)$ is an open subset in $\mathbb{R}^n$. The pair (U, u) is called a *chart* on M.

One of the basic results of algebraic topology, called 'invariance of domain', conjectured by Dedekind and proved by Brouwer in 1911, says that the number n is locally constant on M; if n is constant, M is sometimes called a *pure manifold*. We will only consider pure manifolds and consequently we will omit the prefix pure.

A family $(U_\alpha, u_\alpha)_{\alpha \in A}$ of charts on M such that the U_α form a cover of M is called an *atlas*. The mappings

$$u_{\alpha\beta} := u_\alpha \circ u_\beta^{-1} : u_\beta(U_{\alpha\beta}) \to u_\alpha(U_{\alpha\beta})$$

are called the chart changings for the atlas (U_α), where we use the notation $U_{\alpha\beta} := U_\alpha \cap U_\beta$.

An atlas $(U_\alpha, u_\alpha)_{\alpha \in A}$ for a manifold M is said to be a C^k*-atlas*, if all chart changings $u_{\alpha\beta} : u_\beta(U_{\alpha\beta}) \to u_\alpha(U_{\alpha\beta})$ are differentiable of class C^k. Two C^k-atlases are called C^k*-equivalent* if their union is again a C^k-atlas for M. An equivalence class of C^k-atlases is called a C^k*-structure* on M.

From differential topology we know that if M has a C^1-structure, then it also has a C^1-equivalent C^∞-structure and even a C^1-equivalent C^ω-structure, where C^ω is shorthand for real analytic; see [**84**].

By a C^k-manifold M we mean a topological manifold together with a C^k-structure and a chart on M will be a chart belonging to some atlas of the C^k-structure.

But there are topological manifolds which do not admit differentiable structures. For example, every 4-dimensional manifold is smooth off some point, but there are such which are not smooth; see [**195**], [**62**]. There are also topological manifolds which admit several inequivalent smooth structures. The spheres from dimension 7 on have finitely many; see [**156**]. But the most surprising result is that on $\mathbb{R}^4$ there are uncountably many pairwise inequivalent (exotic) differentiable structures. This follows from the results of [**42**] and [**62**]; see [**78**] for an overview.

Note that for a Hausdorff C^∞-manifold in a more general sense the following properties are equivalent:

(1) It is paracompact.

(2) It is metrizable.

(3) It admits a Riemann metric.

(4) Each connected component is separable.

In this book a manifold will usually mean a C^∞-manifold, and smooth is used synonymously for C^∞ — it will be Hausdorff, separable, finite-dimensional, to state it precisely.

Note finally that any manifold M admits a finite atlas consisting of $\dim M + 1$ (not connected) charts. This is a consequence of topological dimension theory [**168**]; a proof for manifolds may be found in [**80**, I].

1.2. Example: Spheres. We consider the space $\mathbb{R}^{n+1}$, equipped with the standard inner product $\langle x, y\rangle = \sum x^i y^i$. The *n-sphere* S^n is then the subset $\{x \in \mathbb{R}^{n+1} : \langle x, x\rangle = 1\}$. Since $f(x) = \langle x, x\rangle$, $f : \mathbb{R}^{n+1} \to \mathbb{R}$, satisfies $df(x)y = 2\langle x, y\rangle$, it is of rank 1 off 0 and by (1.12) the sphere S^n is a submanifold of $\mathbb{R}^{n+1}$.

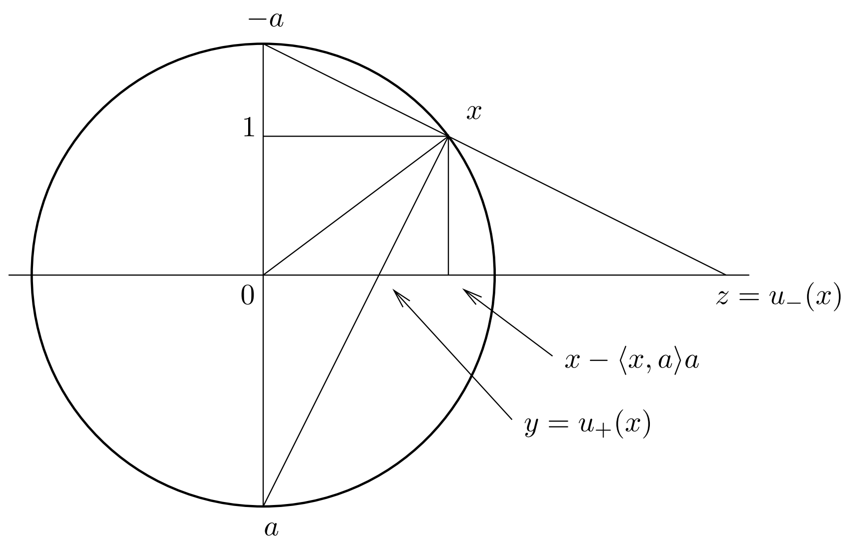
In order to get some feeling for the sphere, we will describe an explicit atlas for S^n, the *stereographic atlas*. Choose $a \in S^n$ ('south pole'). Let

$$U_+ := S^n \setminus \{a\}, \qquad u_+ : U_+ \to \{a\}^\perp, \qquad u_+(x) = \tfrac{x - \langle x,a\rangle a}{1 - \langle x,a\rangle},$$
$$U_- := S^n \setminus \{-a\}, \qquad u_- : U_- \to \{a\}^\perp, \qquad u_-(x) = \tfrac{x - \langle x,a\rangle a}{1 + \langle x,a\rangle}.$$

From the following drawing in the 2-plane through 0, x, and a it is easily seen that u_+ is the usual stereographic projection. We also get

$$u_+^{-1}(y) = \tfrac{|y|^2-1}{|y|^2+1}a + \tfrac{2}{|y|^2+1}y \qquad \text{for } y \in \{a\}^\perp \setminus \{0\}$$

and $(u_- \circ u_+^{-1})(y) = \frac{y}{|y|^2}$. The latter equation can directly be seen from the drawing using the intercept theorem.



1.3. Smooth mappings. A mapping $f : M \to N$ between manifolds is said to be C^k if for each $x \in M$ and one (equivalently: any) chart (V, v) on N with $f(x) \in V$ there is a chart (U, u) on M with $x \in U$, $f(U) \subseteq V$, and $v \circ f \circ u^{-1}$ is C^k. We will denote by $C^k(M, N)$ the space of all C^k-mappings from M to N.

A C^k-mapping $f : M \to N$ is called a C^k-*diffeomorphism* if $f^{-1} : N \to M$ exists and is also C^k. Two manifolds are called *diffeomorphic* if there exists a diffeomorphism between them. From differential topology (see [**84**]) we know that if there is a C^1-diffeomorphism between M and N, then there is also a C^∞-diffeomorphism.

There are manifolds which are homeomorphic but not diffeomorphic: On $\mathbb{R}^4$ there are uncountably many pairwise nondiffeomorphic differentiable structures; on every other $\mathbb{R}^n$ the differentiable structure is unique. There are finitely many different differentiable structures on the spheres S^n for $n \geq 7$.

A mapping $f : M \to N$ between manifolds of the same dimension is called a *local diffeomorphism* if each $x \in M$ has an open neighborhood U such that $f|U : U \to f(U) \subset N$ is a diffeomorphism. Note that a local diffeomorphism need not be surjective.

1.4. Smooth functions. The set of smooth real valued functions on a manifold M will be denoted by $C^\infty(M)$, in order to distinguish it clearly from spaces of sections which will appear later. The space $C^\infty(M)$ is a real commutative algebra.

The *support* of a smooth function f is the closure of the set where it does not vanish, $\operatorname{supp}(f) = \overline{\{x \in M : f(x) \neq 0\}}$. The *zero set* of f is the set where f vanishes, $Z(f) = \{x \in M : f(x) = 0\}$.

1.5. Theorem. *Any (separable, metrizable, smooth) manifold admits smooth partitions of unity: Let $(U_\alpha)_{\alpha\in A}$ be an open cover of M.*

Then there is a family $(\varphi_\alpha)_{\alpha\in A}$ of smooth functions on M, such that:

(1) $\varphi_\alpha(x) \geq 0$ *for all $x \in M$ and all $\alpha \in A$.*
(2) $\operatorname{supp}(\varphi_\alpha) \subset U_\alpha$ *for all $\alpha \in A$.*
(3) *$(supp(\varphi_\alpha))_{\alpha\in A}$ is a locally finite family (so each $x \in M$ has an open neighborhood which meets only finitely many* $\operatorname{supp}(\varphi_\alpha)$*).*
(4) $\sum_\alpha \varphi_\alpha = 1$ *(locally this is a finite sum).*

Proof. Any (separable, metrizable) manifold is a *'Lindelöf space'*, i.e., each open cover admits a countable subcover. This can be seen as follows:

Let $\mathcal{U}$ be an open cover of M. Since M is separable, there is a countable dense subset S in M. Choose a metric on M. For each $U \in \mathcal{U}$ and each $x \in U$ there is a $y \in S$ and $n \in \mathbb{N}$ such that the ball $B_{1/n}(y)$ with respect to that metric with center y and radius $\frac{1}{n}$ contains x and is contained in U. But there are only countably many of these balls; for each of them we choose an open set $U \in \mathcal{U}$ containing it. This is then a countable subcover of $\mathcal{U}$.

Now let $(U_\alpha)_{\alpha\in A}$ be the given cover. Let us fix first α and $x \in U_\alpha$. We choose a chart (U, u) centered at x (i.e., $u(x) = 0$) and $\varepsilon > 0$ such that $\varepsilon\mathbb{D}^n \subset u(U \cap U_\alpha)$, where $\mathbb{D}^n = \{y \in \mathbb{R}^n : |y| \leq 1\}$ is the closed unit ball. Let

$$h(t) := \begin{cases} e^{-1/t} & \text{for } t > 0, \\ 0 & \text{for } t \leq 0, \end{cases}$$

a smooth function on $\mathbb{R}$. Then

$$f_{\alpha,x}(z) := \begin{cases} h(\varepsilon^2 - |u(z)|^2) & \text{for } z \in U, \\ 0 & \text{for } z \notin U \end{cases}$$

is a nonnegative smooth function on M with support in U_α which is positive at x.

We choose such a function $f_{\alpha,x}$ for each α and $x \in U_\alpha$. The interiors of the supports of these smooth functions form an open cover of M which refines

(U_α), so by the argument at the beginning of the proof there is a countable subcover with corresponding functions $f_1, f_2, \dots$. Let

$$W_n = \{x \in M : f_n(x) > 0 \text{ and } f_i(x) < \tfrac{1}{n} \text{ for } 1 \le i < n\},$$

and denote by $\overline{W}_n$ the closure. Then $(W_n)_n$ is an open cover. We claim that $(\overline{W}_n)_n$ is locally finite: Let $x \in M$. Then there is a smallest n such that $x \in W_n$. Let $V := \{y \in M : f_n(y) > \frac{1}{2} f_n(x)\}$. If $y \in V \cap \overline{W}_k$, then we have $f_n(y) > \frac{1}{2} f_n(x)$ and $f_i(y) \le \frac{1}{k}$ for $i < k$, which is possible for finitely many k only.

Consider the nonnegative smooth function

$$g_n(x) = h(f_n(x)) h(\tfrac{1}{n} - f_1(x)) \dots h(\tfrac{1}{n} - f_{n-1}(x)), \quad n \in \mathbb{N}.$$

Then obviously $\operatorname{supp}(g_n) = \overline{W}_n$. So $g := \sum_n g_n$ is smooth, since it is locally only a finite sum, and everywhere positive; thus $(g_n/g)_{n \in \mathbb{N}}$ is a smooth partition of unity on M. Since $\operatorname{supp}(g_n) = \overline{W}_n$ is contained in some $U_{\alpha(n)}$, we may put $\varphi_\alpha = \sum_{\{n : \alpha(n) = \alpha\}} \frac{g_n}{g}$ to get the required partition of unity which is subordinated to $(U_\alpha)_{\alpha \in A}$. $\square$

1.6. Germs. Let M and N be manifolds and $x \in M$. We consider all smooth mappings $f : U_f \to N$, where U_f is some open neighborhood of x in M, and we put $f \sim_x g$ if there is some open neighborhood V of x with $f|V = g|V$. This is an equivalence relation on the set of mappings considered. The equivalence class of a mapping f is called the *germ of f at x*, sometimes denoted by $\operatorname{germ}_x f$. The set of all these germs is denoted by $C^\infty_x(M, N)$.

Note that for a germs at x of a smooth mapping only the value at x is defined. We may also consider composition of germs: $\operatorname{germ}_{f(x)} g \circ \operatorname{germ}_x f := \operatorname{germ}_x(g \circ f)$.

If $N = \mathbb{R}$, we may add and multiply germs of smooth functions, so we get the real commutative algebra $C^\infty_x(M, \mathbb{R})$ of germs of smooth functions at x. This construction works also for other types of functions like real analytic or holomorphic ones if M has a real analytic or complex structure.

Using smooth partitions of unity (1.4) it is easily seen that each germ of a smooth function has a representative which is defined on the whole of M. For germs of real analytic or holomorphic functions this is not true. So $C^\infty_x(M, \mathbb{R})$ is the quotient of the algebra $C^\infty(M)$ by the ideal of all smooth functions $f : M \to \mathbb{R}$ which vanish on some neighborhood (depending on f) of x.

1.7. The tangent space of $\mathbb{R}^n$. Let $a \in \mathbb{R}^n$. A *tangent vector* with foot point a is simply a pair (a, X) with $X \in \mathbb{R}^n$, also denoted by X_a. It induces a *derivation* $X_a : C^\infty(\mathbb{R}^n) \to \mathbb{R}$ by $X_a(f) = df(a)(X_a)$. The value depends only on the germ of f at a and we have $X_a(f\cdot g) = X_a(f)\cdot g(a) + f(a)\cdot X_a(g)$ (the derivation property).

If conversely $D : C^\infty(\mathbb{R}^n) \to \mathbb{R}$ is linear and satisfies

$$D(f\cdot g) = D(f)\cdot g(a) + f(a)\cdot D(g)$$

(a derivation at a), then D is given by the action of a tangent vector with foot point a. This can be seen as follows. For $f \in C^\infty(\mathbb{R}^n)$ we have

$$\begin{aligned} f(x) &= f(a) + \int_0^1 \tfrac{d}{dt} f(a + t(x-a))dt \\ &= f(a) + \sum_{i=1}^n \int_0^1 \tfrac{\partial f}{\partial x^i}(a + t(x-a))dt\ (x^i - a^i) \\ &= f(a) + \sum_{i=1}^n h_i(x)(x^i - a^i). \end{aligned}$$

On the constant function 1 the derivation gives $D(1) = D(1\cdot 1) = 2D(1)$, so $D(\text{constant}) = 0$. Therefore,

$$\begin{aligned} D(f) &= D\Big(f(a) + \sum_{i=1}^n h_i(x^i - a^i)\Big) \\ &= 0 + \sum_{i=1}^n D(h_i)(a^i - a^i) + \sum_{i=1}^n h_i(a)(D(x^i) - 0) \\ &= \sum_{i=1}^n \tfrac{\partial f}{\partial x^i}(a) D(x^i), \end{aligned}$$

where x^i is the i-th coordinate function on $\mathbb{R}^n$. So we have

$$D(f) = \sum_{i=1}^n D(x^i)\tfrac{\partial}{\partial x^i}|_a(f), \qquad D = \sum_{i=1}^n D(x^i)\tfrac{\partial}{\partial x^i}|_a.$$

Thus D is induced by the tangent vector $(a, \sum_{i=1}^n D(x^i)e_i)$, where (e_i) is the standard basis of $\mathbb{R}^n$.

1.8. The tangent space of a manifold. Let M be a manifold and let $x \in M$ and $\dim M = n$. Let T_xM be the vector space of all derivations at x of $C^\infty_x(M,\mathbb{R})$, the algebra of germs of smooth functions on M at x. Using (1.5), it may easily be seen that a derivation of $C^\infty(M)$ at x factors to a derivation of $C^\infty_x(M,\mathbb{R})$.

So T_xM consists of all linear mappings $X_x : C^\infty(M) \to \mathbb{R}$ with the property $X_x(f\cdot g) = X_x(f)\cdot g(x) + f(x)\cdot X_x(g)$. The space T_xM is called the *tangent space of M at x*.

If (U,u) is a chart on M with $x \in U$, then $u^* : f \mapsto f \circ u$ induces an isomorphism of algebras $C^\infty_{u(x)}(\mathbb{R}^n,\mathbb{R}) \cong C^\infty_x(M,\mathbb{R})$, and thus also an isomorphism $T_x u : T_x M \to T_{u(x)}\mathbb{R}^n$, given by $(T_x u.X_x)(f) = X_x(f \circ u)$. So $T_x M$ is an n-dimensional vector space.

We will use the following notation: $u = (u^1,\dots,u^n)$, so u^i denotes the i-th coordinate function on U, and

$$\tfrac{\partial}{\partial u^i}|_x := (T_x u)^{-1}(\tfrac{\partial}{\partial x^i}|_{u(x)}) = (T_x u)^{-1}(u(x), e_i).$$

So $\frac{\partial}{\partial u^i}|_x \in T_x M$ is the derivation given by

$$\tfrac{\partial}{\partial u^i}|_x(f) = \frac{\partial(f \circ u^{-1})}{\partial x^i}(u(x)).$$

From (1.7) we have now

$$\begin{aligned}
T_x u.X_x &= \sum_{i=1}^n (T_x u.X_x)(x^i)\tfrac{\partial}{\partial x^i}|_{u(x)} = \sum_{i=1}^n X_x(x^i \circ u)\tfrac{\partial}{\partial x^i}|_{u(x)}\\
&= \sum_{i=1}^n X_x(u^i)\tfrac{\partial}{\partial x^i}|_{u(x)},\\
X_x &= (T_x u)^{-1}.T_x u.X_x = \sum_{i=1}^n X_x(u^i)\tfrac{\partial}{\partial u^i}|_x.
\end{aligned}$$

1.9. The tangent bundle. For a manifold M of dimension n we put $TM := \bigsqcup_{x\in M} T_x M$, the disjoint union of all tangent spaces. This is a family of vector spaces parameterized by M, with projection $\pi_M : TM \to M$ given by $\pi_M(T_x M) = x$.

For any chart (U_α, u_α) of M consider the chart $(\pi_M^{-1}(U_\alpha), Tu_\alpha)$ on TM, where $Tu_\alpha : \pi_M^{-1}(U_\alpha) \to u_\alpha(U_\alpha) \times \mathbb{R}^n$ is given by

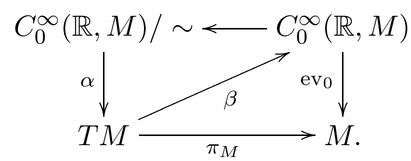
$$Tu_\alpha.X = (u_\alpha(\pi_M(X)), T_{\pi_M(X)}u_\alpha.X).$$

Then the chart changings look as follows:

$$\begin{aligned}
Tu_\beta \circ (Tu_\alpha)^{-1} : Tu_\alpha(\pi_M^{-1}(U_{\alpha\beta})) &= u_\alpha(U_{\alpha\beta}) \times \mathbb{R}^n \to\\
&\to u_\beta(U_{\alpha\beta}) \times \mathbb{R}^n = Tu_\beta(\pi_M^{-1}(U_{\alpha\beta})),\\
((Tu_\beta \circ (Tu_\alpha)^{-1})(y,Y))(f) &= ((Tu_\alpha)^{-1}(y,Y))(f \circ u_\beta)\\
&= (y,Y)(f \circ u_\beta \circ u_\alpha^{-1}) = d(f \circ u_\beta \circ u_\alpha^{-1})(y).Y\\
&= df(u_\beta \circ u_\alpha^{-1}(y)).d(u_\beta \circ u_\alpha^{-1})(y).Y\\
&= (u_\beta \circ u_\alpha^{-1}(y), d(u_\beta \circ u_\alpha^{-1})(y).Y)(f).
\end{aligned}$$

So the chart changings are smooth. We choose the topology on TM in such a way that all Tu_α become homeomorphisms. This is a Hausdorff topology, since $X, Y \in TM$ may be separated in M if $\pi(X) \neq \pi(Y)$; and they may be separated in one chart if $\pi(X) = \pi(Y)$. So TM is again a smooth manifold in a canonical way; the triple (TM, π_M, M) is called the *tangent bundle* of the manifold M.

1.10. Kinematic definition of the tangent space. Let $C_0^\infty(\mathbb{R}, M)$ denote the space of germs at 0 of smooth curves $\mathbb{R} \to M$. We put the following equivalence relation on $C_0^\infty(\mathbb{R}, M)$: the germ of c is equivalent to the germ of e if and only if $c(0) = e(0)$ and in one (equivalently: each) chart (U, u) with $c(0) = e(0) \in U$ we have $\frac{d}{dt}|_0(u \circ c)(t) = \frac{d}{dt}|_0(u \circ e)(t)$. The equivalence classes are also called velocity vectors of curves in M. We have the following diagram of mappings where $\alpha(c)(\text{germ}_{c(0)} f) = \frac{d}{dt}|_0 f(c(t))$ and $\beta : TM \to C_0^\infty(\mathbb{R}, M)$ is given by: $\beta((Tu)^{-1}(y, Y))$ is the germ at 0 of $t \mapsto u^{-1}(y + tY)$. So TM is canonically identified with the set of all possible velocity vectors of curves in M:



1.11. Tangent mappings. Let $f : M \to N$ be a smooth mapping between manifolds. Then f induces a linear mapping $T_x f : T_x M \to T_{f(x)} N$ for each $x \in M$ by $(T_x f.X_x)(h) = X_x(h \circ f)$ for $h \in C^\infty_{f(x)}(N, \mathbb{R})$. This mapping is well defined and linear since $f^* : C^\infty_{f(x)}(N, \mathbb{R}) \to C^\infty_x(M, \mathbb{R})$, given by $h \mapsto h \circ f$, is linear and an algebra homomorphism, and $T_x f$ is its adjoint, restricted to the subspace of derivations.

If (U, u) is a chart around x and (V, v) is one around $f(x)$, then

$$\begin{aligned}
(T_x f.\tfrac{\partial}{\partial u^i}|_x)(v^j) &= \tfrac{\partial}{\partial u^i}|_x(v^j \circ f) = \tfrac{\partial}{\partial x^i}(v^j \circ f \circ u^{-1})(u(x)),\\
T_x f.\tfrac{\partial}{\partial u^i}|_x &= \textstyle\sum_j (T_x f.\tfrac{\partial}{\partial u^i}|_x)(v^j) \tfrac{\partial}{\partial v^j}|_{f(x)} \qquad \text{by (1.8)}\\
&= \textstyle\sum_j \tfrac{\partial(v^j \circ f \circ u^{-1})}{\partial x^i}(u(x)) \tfrac{\partial}{\partial v^j}|_{f(x)}.
\end{aligned}$$

So the matrix of $T_x f : T_x M \to T_{f(x)} N$ in the bases $(\frac{\partial}{\partial u^i}|_x)$ and $(\frac{\partial}{\partial v^j}|_{f(x)})$ is just the Jacobi matrix $d(v \circ f \circ u^{-1})(u(x))$ of the mapping $v \circ f \circ u^{-1}$ at $u(x)$, so $T_{f(x)} v \circ T_x f \circ (T_x u)^{-1} = d(v \circ f \circ u^{-1})(u(x))$.

Let us denote by $Tf : TM \to TN$ the total mapping which is given by $Tf|T_x M := T_x f$. Then the composition

$$\begin{aligned}
Tv \circ Tf \circ (Tu)^{-1} &: u(U) \times \mathbb{R}^m \to v(V) \times \mathbb{R}^n,\\
(y, Y) &\mapsto ((v \circ f \circ u^{-1})(y), d(v \circ f \circ u^{-1})(y)Y),
\end{aligned}$$

is smooth; thus $Tf : TM \to TN$ is again smooth.

If $f : M \to N$ and $g : N \to P$ are smooth, then we have $T(g \circ f) = Tg \circ Tf$. This is a direct consequence of $(g \circ f)^* = f^* \circ g^*$, and it is the global version of the chain rule. Furthermore we have $T(Id_M) = Id_{TM}$.

If $f \in C^\infty(M)$, then $Tf : TM \to T\mathbb{R} = \mathbb{R} \times \mathbb{R}$. We define the *differential* of f by $df := \mathrm{pr}_2 \circ Tf : TM \to \mathbb{R}$. Let t denote the identity function on $\mathbb{R}$. Then $(Tf.X_x)(t) = X_x(t \circ f) = X_x(f)$, so we have $df(X_x) = X_x(f)$.

1.12. Submanifolds. A subset N of a manifold M is called a *submanifold* if for each $x \in N$ there is a chart (U, u) of M such that $u(U \cap N) = u(U) \cap (\mathbb{R}^k \times 0)$, where $\mathbb{R}^k \times 0 \hookrightarrow \mathbb{R}^k \times \mathbb{R}^{n-k} = \mathbb{R}^n$. Then clearly N is itself a manifold with $(U \cap N, u|(U \cap N))$ as charts, where (U, u) runs through all *submanifold charts* as above.

1.13. Let $f : \mathbb{R}^n \to \mathbb{R}^q$ be smooth. A point $x \in \mathbb{R}^q$ is called a *regular value* of f if the rank of f (more exactly: the rank of its derivative) is q at each point y of $f^{-1}(x)$. In this case, $f^{-1}(x)$ is a submanifold of $\mathbb{R}^n$ of dimension $n - q$ (or empty). This is an immediate consequence of the implicit function theorem, as follows: Let $x = 0 \in \mathbb{R}^q$. Permute the coordinates $(x^1, \dots, x^n)$ on $\mathbb{R}^n$ such that the Jacobi matrix

$$df(y) = \left(\left(\frac{\partial f^i}{\partial x^j}(y) \right)^{1 \le i \le q}_{1 \le j \le q} \middle| \left(\frac{\partial f^i}{\partial x^j}(y) \right)^{1 \le i \le q}_{q+1 \le j \le n} \right)$$

has the left hand part invertible. Then $u := (f, \mathrm{pr}_{n-q}) : \mathbb{R}^n \to \mathbb{R}^q \times \mathbb{R}^{n-q}$ has invertible differential at y, so (U, u) is a chart at any $y \in f^{-1}(0)$, and we have $f \circ u^{-1}(z^1, \dots, z^n) = (z^1, \dots, z^q)$, so $u(f^{-1}(0)) = u(U) \cap (0 \times \mathbb{R}^{n-q})$ as required.

Constant rank theorem ([41, I 10.3.1]). *Let $f : W \to \mathbb{R}^q$ be a smooth mapping, where W is an open subset of $\mathbb{R}^n$. If the derivative $df(x)$ has constant rank k for each $x \in W$, then for each $a \in W$ there are charts (U, u) of W centered at a and (V, v) of $\mathbb{R}^q$ centered at $f(a)$ such that $v \circ f \circ u^{-1} : u(U) \to v(V)$ has the following form:*

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_k, 0, \dots, 0).$$

So $f^{-1}(b)$ is a submanifold of W of dimension $n - k$ for each $b \in f(W)$.

Proof. We will use the inverse function theorem several times. The derivative $df(a)$ has rank $k \le n, q$; without loss we may assume that the upper left $(k \times k)$-submatrix of $df(a)$ is invertible. Moreover, let $a = 0$ and $f(a) = 0$. We consider $u : W \to \mathbb{R}^n$, $u(x^1, \dots, x^n) := (f^1(x), \dots, f^k(x), x^{k+1}, \dots, x^n)$. Then

$$du = \begin{pmatrix} (\frac{\partial f^i}{\partial z^j})^{1 \le i \le k}_{1 \le j \le k} & (\frac{\partial f^i}{\partial z^j})^{1 \le i \le k}_{k+1 \le j \le n} \\ 0 & \mathbb{I}_{\mathbb{R}^{n-k}} \end{pmatrix}$$

is invertible, so u is a diffeomorphism $U_1 \to U_2$ for suitable open neighborhoods of 0 in $\mathbb{R}^n$. Consider $g = f \circ u^{-1} : U_2 \to \mathbb{R}^q$. Then we have

$$\begin{aligned} g(z_1, \dots, z_n) &= (z_1, \dots, z_k, g_{k+1}(z), \dots, g_q(z)), \\ dg(z) &= \begin{pmatrix} \mathbb{I}_{\mathbb{R}^k} & 0 \\ * & (\frac{\partial g^i}{\partial z^j})_{k+1 \le j \le n}^{k+1 \le i \le q} \end{pmatrix}, \\ \operatorname{rank}(dg(z)) &= \operatorname{rank}\big(d(f \circ u^{-1})(z)\big) \\ &= \operatorname{rank}\big(df(u^{-1}(z)).du^{-1}(z)\big) = \operatorname{rank}(df(z)) = k. \end{aligned}$$

Therefore, $$\frac{\partial g^i}{\partial z^j}(z) = 0 \qquad \text{for } k+1 \le i \le q \text{ and } k+1 \le j \le n;$$

$$g^i(z^1, \dots, z^n) = g^i(z^1, \dots, z^k, 0, \dots, 0) \quad \text{for } k+1 \le i \le q.$$

Let $v : U_3 \to \mathbb{R}^q$, where $U_3 = \{y \in \mathbb{R}^q : (y^1, \dots, y^k, 0, \dots, 0) \in U_2 \subset \mathbb{R}^n\}$, be given by

$$v \begin{pmatrix} y^1 \\ \vdots \\ y^q \end{pmatrix} = \begin{pmatrix} y^1 \\ \vdots \\ y^k \\ y^{k+1} - g^{k+1}(y^1, \dots, y^k, 0, \dots, 0) \\ \vdots \\ y^q - g^q(y^1, \dots, y^k, 0, \dots, 0) \end{pmatrix} = \begin{pmatrix} y^1 \\ \vdots \\ y^k \\ y^{k+1} - g^{k+1}(\bar{y}) \\ \vdots \\ y^q - g^q(\bar{y}) \end{pmatrix},$$

where $\bar{y} = (y^1, \dots, y^q, 0, \dots, 0) \in \mathbb{R}^n$ if $q < n$ and $\bar{y} = (y^1, \dots, y^n)$ if $q \ge n$. We have $v(0) = 0$, and

$$dv = \begin{pmatrix} \mathbb{I}_{\mathbb{R}^k} & 0 \\ * & \mathbb{I}_{\mathbb{R}^{q-k}} \end{pmatrix}$$

is invertible; thus $v : V \to \mathbb{R}^q$ is a chart for a suitable neighborhood of 0. Now let $U := f^{-1}(V) \cup U_1$. Then $v \circ f \circ u^{-1} = v \circ g : \mathbb{R}^n \supseteq u(U) \to v(V) \subseteq \mathbb{R}^q$ looks as follows:

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \xrightarrow{g} \begin{pmatrix} x^1 \\ \vdots \\ x^k \\ g^{k+1}(x) \\ \vdots \\ g^q(x) \end{pmatrix} \xrightarrow{v} \begin{pmatrix} x^1 \\ \vdots \\ x^k \\ g^{k+1}(x) - g^{k+1}(x) \\ \vdots \\ g^q(x) - g^q(x) \end{pmatrix} = \begin{pmatrix} x^1 \\ \vdots \\ x^k \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad \square$$

Corollary. *Let $f : M \to N$ be C^∞ with $T_x f$ of constant rank k for all $x \in M$.*

Then for each $b \in f(M)$ the set $f^{-1}(b) \subset M$ is a submanifold of M of dimension $\dim M - k$. $\square$

1.14. Products. Let M and N be smooth manifolds described by smooth atlases $(U_\alpha, u_\alpha)_{\alpha\in A}$ and $(V_\beta, v_\beta)_{\beta\in B}$, respectively. Then the family $(U_\alpha \times V_\beta, u_\alpha \times v_\beta : U_\alpha \times V_\beta \to \mathbb{R}^m \times \mathbb{R}^n)_{(\alpha,\beta)\in A\times B}$ is a smooth atlas for the cartesian product $M \times N$. Clearly the projections

$$M \xleftarrow{\mathrm{pr}_1} M \times N \xrightarrow{\mathrm{pr}_2} N$$

are also smooth. The *product* $(M \times N, \mathrm{pr}_1, \mathrm{pr}_2)$ has the following universal property:

> For any smooth manifold P and smooth mappings $f : P \to M$ and $g : P \to N$ the mapping
>
> $$(f, g) : P \to M \times N, \quad (f, g)(x) = (f(x), g(x)),$$
>
> is the unique smooth mapping with $\mathrm{pr}_1 \circ (f, g) = f$ and $\mathrm{pr}_2 \circ (f, g) = g$.

From the construction of the tangent bundle in (1.9) it is immediately clear that

$$TM \xleftarrow{T(\mathrm{pr}_1)} T(M \times N) \xrightarrow{T(\mathrm{pr}_2)} TN$$

is again a product, so that $T(M \times N) = TM \times TN$ in a canonical way.

Clearly we can form products of finitely many manifolds.

1.15. Theorem. *Let M be a connected manifold and suppose that $f : M \to M$ is smooth with $f \circ f = f$. Then the image $f(M)$ of f is a submanifold of M.*

This result can also be expressed as: 'smooth retracts' of manifolds are manifolds. If we do not suppose that M is connected, then $f(M)$ will not be a pure manifold in general; it will have different dimensions in different connected components.

Proof. We claim that there is an open neighborhood U of $f(M)$ in M such that the rank of $T_y f$ is constant for $y \in U$. Then by theorem (1.13) the result follows.

For $x \in f(M)$ we have $T_x f \circ T_x f = T_x f$; thus $\operatorname{im} T_x f = \ker(Id - T_x f)$ and $\operatorname{rank} T_x f + \operatorname{rank}(Id - T_x f) = \dim M$. Since $\operatorname{rank} T_x f$ and $\operatorname{rank}(Id - T_x f)$ cannot fall locally, $\operatorname{rank} T_x f$ is locally constant for $x \in f(M)$, and since $f(M)$ is connected, $\operatorname{rank} T_x f = r$ for all $x \in f(M)$.

But then for each $x \in f(M)$ there is an open neighborhood U_x in M with $\operatorname{rank} T_y f \ge r$ for all $y \in U_x$. On the other hand

$$\operatorname{rank} T_y f = \operatorname{rank} T_y (f \circ f) = \operatorname{rank} T_{f(y)} f \circ T_y f \le \operatorname{rank} T_{f(y)} f = r$$

since $f(y) \in f(M)$.

So the neighborhood we need is given by $U = \bigcup_{x\in f(M)} U_x$. $\square$

1.16. Corollary. (1) *The (separable) connected smooth manifolds are exactly the smooth retracts of connected open subsets of* $\mathbb{R}^n$*'s.*

(2) *A smooth mapping* $f : M \to N$ *is an embedding of a submanifold if and only if there is an open neighborhood* U *of* $f(M)$ *in* N *and a smooth mapping* $r : U \to M$ *with* $r \circ f = Id_M$*.*

Proof. Any manifold M may be embedded into some $\mathbb{R}^n$; see (1.19) below. Then there exists a tubular neighborhood of M in R^n (see later or [**84**, pp. 109–118]), and M is clearly a retract of such a tubular neighborhood. The converse follows from (1.15).

For the second assertion we repeat the argument for N instead of $\mathbb{R}^n$. □

1.17. Sets of Lebesque measure 0 in manifolds. An *m-cube of width $w > 0$ in $\mathbb{R}^m$* is a set of the form $C = [x_1, x_1 + w] \times \ldots \times [x_m, x_m + w]$. The measure $\mu(C)$ is then $\mu(C) = w^n$. A subset $S \subset \mathbb{R}^m$ is called a *set of (Lebesque) measure* 0 if for each $\varepsilon > 0$ these are at most countably many m-cubes C_i with $S \subset \bigcup_{i=0}^{\infty} C_i$ and $\sum_{i=0}^{\infty} \mu(C_i) < \varepsilon$. Obviously, a countable union of sets of Lebesque measure 0 is again of measure 0.

Lemma. *Let* $U \subset \mathbb{R}^m$ *be open and let* $f : U \to \mathbb{R}^m$ *be* C^1*. If* $S \subset U$ *is of measure* 0*, then also* $f(S) \subset \mathbb{R}^m$ *is of measure* 0*.*

Proof. Every point of S belongs to an open ball $B \subset U$ such that the operator norm $\|df(x)\| \le K_B$ for all $x \in B$. Then $|f(x) - f(y)| \le K_B|x - y|$ for all $x, y \in B$. So if $C \subset B$ is an m-cube of width w, then $f(C)$ is contained in an m-cube C' of width $\sqrt{m}K_B w$ and measure $\mu(C') \le m^{m/2}K_B^m \mu(C)$. Now let $S = \bigcup_{j=1}^{\infty} S_j$ where each S_j is a compact subset of a ball B_j as above. It suffices to show that each $f(S_j)$ is of measure 0.

For each $\varepsilon > 0$ there are m-cubes C_i in B_j with $S_j \subset \bigcup_i C_i$ and $\sum_i \mu(C_i) < \varepsilon$. As we saw above, then $f(X_j) \subset \bigcup_i C_i'$ with $\sum_i \mu(C_i') < m^{m/2}K_{B_j}^m \varepsilon$. □

Let M be a smooth (separable) manifold. A subset $S \subset M$ is called a *set of (Lebesque) measure* 0 if for each chart (U, u) of M the set $u(S \cap U)$ is of measure 0 in $\mathbb{R}^m$. By the lemma it suffices that there is some atlas whose charts have this property. Obviously, a countable union of sets of measure 0 in a manifold is again of measure 0.

An m-cube is not of measure 0. Thus a subset of $\mathbb{R}^m$ of measure 0 does not contain any m-cube; hence its interior is empty. Thus a closed set of measure 0 in a manifold is nowhere dense. More generally, let S be a subset of a manifold which is of measure 0 and σ-compact, i.e., a countable union of compact subsets. Then each of the latter is nowhere dense, so S is nowhere dense by the Baire category theorem. The complement of S is *residual*, i.e., it contains the intersection of a countable family of open dense subsets.

The Baire theorem says that a residual subset of a complete metric space is dense.

1.18. Regular values. Let $f : M \to N$ be a smooth mapping between manifolds.

(1) A point $x \in M$ is called a *singular point* of f if $T_x f$ is not surjective, and it is called a *regular point* of f if $T_x f$ is surjective.

(2) A point $y \in N$ is called a *regular value* of f if $T_x f$ is surjective for all $x \in f^{-1}(y)$. If not, y is called a *singular value*. Note that any $y \in N \setminus f(M)$ is a regular value.

Theorem ([166], [196]). *The set of all singular values of a C^k mapping $f : M \to N$ is of Lebesgue measure 0 in N if $k > \max\{0, \dim(M) - \dim(N)\}$.*

So any smooth mapping has regular values.

Proof. We proof this only for smooth mappings. It is sufficient to prove this locally. Thus we consider a smooth mapping $f : U \to \mathbb{R}^n$ where $U \subset \mathbb{R}^m$ is open. If $n > m$, then the result follows from lemma (1.17) above (consider the set $U \times 0 \subset \mathbb{R}^m \times \mathbb{R}^{n-m}$ of measure 0). Thus let $m \geq n$.

Let $\Sigma(f) \subset U$ denote the set of singular points of f. Let $f = (f^1, \dots, f^n)$, and let $\Sigma(f) = \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$ where:

Σ_1 is the set of singular points x such that $Pf(x) = 0$ for all linear differential operators P of order $\leq \frac{m}{n}$.

Σ_2 is the set of singular points x such that $Pf(x) \neq 0$ for some differential operator P of order ≥ 2.

Σ_3 is the set of singular points x such that $\frac{\partial f^i}{x^j}(x) = 0$ for some i, j.

We first show that $f(\Sigma_1)$ has measure 0. Let $\nu = \lceil \frac{m}{n} + 1 \rceil$ be the smallest integer $> m/n$. Then each point of Σ_1 has an open neighborhood $W \subset U$ such that $|f(x) - f(y)| \leq K|x - y|^\nu$ for all $x \in \Sigma_1 \cap W$ and $y \in W$ and for some $K > 0$, by Taylor expansion. We take W to be a cube, of width w. It suffices to prove that $f(\Sigma_1 \cap W)$ has measure 0. We divide W into p^m cubes of width $\frac{w}{p}$; those which meet Si_1 will be denoted by $C_1, \dots, C_q$ for $q \leq p^m$. Each C_k is contained in a ball of radius $\frac{w}{p}\sqrt{m}$ centered at a point of $\Sigma_1 \cap W$. The set $f(C_k)$ is contained in a cube $C'_k \subset \mathbb{R}^n$ of width $2K(\frac{w}{p}\sqrt{m})^\nu$. Then

$$\sum_k \mu^n(C'_k) \leq p^m (2K)^n (\frac{w}{p}\sqrt{m})^{\nu n} = p^{m-\nu n}(2K)^n w^{\nu n} \to 0 \text{ for } p \to \infty,$$

since $m - \nu n < 0$.

Note that $\Sigma(f) = \Sigma_1$ if $n = m = 1$. So the theorem is proved in this case. We proceed by induction on m. So let $m > 1$ and assume that the theorem is true for each smooth map $P \to Q$ where $\dim(P) < m$.

We prove that $f(\Sigma_2 \setminus \Sigma_3)$ has measure 0. For each $x \in \Sigma_2 \setminus \Sigma_3$ there is a linear differential operator P such that $Pf(x) = 0$ and $\frac{\partial f^i}{\partial x^j}(x) \neq 0$ for some i, j. Let W be the set of all such points, for fixed P, i, j. It suffices to show that $f(W)$ has measure 0. By assumption, $0 \in \mathbb{R}$ is a regular value for the function $Pf^i : W \to \mathbb{R}$. Therefore W is a smooth submanifold of dimension $m-1$ in $\mathbb{R}^m$. Clearly, $\Sigma(f) \cap W$ is contained in the set of all singular points of $f|W : W \to \mathbb{R}^n$, and by induction we get that $f((\Sigma_2 \setminus \Sigma_3) \cap W) \subset f(\Sigma(f) \cap W) \subset f(\Sigma(f|W))$ has measure 0.

It remains to prove that $f(\Sigma_3)$ has measure 0. Every point of Σ_3 has an open neighborhood $W \subset U$ on which $\frac{\partial f^i}{\partial x^j} \neq 0$ for some i, j. By shrinking W if necessary and applying diffeomorphisms, we may assume that

$$\mathbb{R}^{m-1} \times \mathbb{R} \supseteq W_1 \times W_2 = W \xrightarrow{f} \mathbb{R}^{n-1} \times \mathbb{R}, \qquad (y,t) \mapsto (g(y,t), t).$$

Clearly, (y, t) is a critical point for f if and only if y is a critical point for $g(\quad, t)$. Thus $\Sigma(f) \cap W = \bigcup_{t \in W_2} (\Sigma(g(\quad, t)) \times \{t\})$. Since $\dim(W_1) = m-1$, by induction we get that $\mu^{n-1}(g(\Sigma(g(\quad, t), t))) = 0$, where μ^{n-1} is the Lebesque measure in $\mathbb{R}^{n-1}$. By Fubini's theorem we get

$$\begin{aligned} \mu^n\Big(\bigcup_{t \in W_2} (\Sigma(g(\quad, t)) \times \{t\})\Big) &= \int_{W_2} \mu^{n-1}(g(\Sigma(g(\quad, t), t)))\, dt \\ &= \int_{W_2} 0\, dt = 0. \quad \square \end{aligned}$$

1.19. Embeddings into $\mathbb{R}^n$'s. Let M be a smooth manifold of dimension m. Then M can be embedded into $\mathbb{R}^n$ if

(1) $n = 2m + 1$ (this is due to [**228**]; see also [**84**, p. 55] or [**26**, p. 73]).
(2) $n = 2m$ (see [**228**]).
(3) Conjecture (still unproved): The minimal n is $n = 2m - \alpha(m) + 1$, where $\alpha(m)$ is the number of 1's in the dyadic expansion of m.

There exists an immersion (see section (2)) $M \to \mathbb{R}^n$ if

(4) $n = 2m$ (see [**84**]).
(5) $n = 2m - 1$ (see [**228**]).
(6) Conjecture: The minimal n is $n = 2m - \alpha(m)$. The article [**34**] claims to have proven this. The proof is believed to be incomplete.

Examples and Exercises

1.20. Discuss the following submanifolds of $\mathbb{R}^n$; in particular make drawings of them:

The unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : \langle x, x\rangle = 1\} \subset \mathbb{R}^n$.

The *ellipsoid* $\{x \in \mathbb{R}^n : f(x) := \sum_{i=1}^n \frac{x_i^2}{a_i^2} = 1\}$, $a_i \neq 0$, with principal axis $a_1, \dots, a_n$.

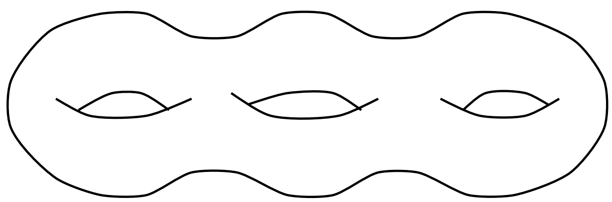
The *hyperboloid* $\{x \in \mathbb{R}^n : f(x) := \sum_{i=1}^n \varepsilon_i \frac{x_i^2}{a_i^2} = 1\}$, $\varepsilon_i = \pm 1$, $a_i \neq 0$, with principal axis a_i and index $= \sum \varepsilon_i$.

The *saddle* $\{x \in \mathbb{R}^3 : x_3 = x_1 x_2\}$.

The *torus*: the rotation surface generated by rotation of $(y-R)^2 + z^2 = r^2$, $0 < r < R$, with center the z–axis, i.e.,

$$\{(x, y, z) : (\sqrt{x^2 + y^2} - R)^2 + z^2 = r^2\}.$$

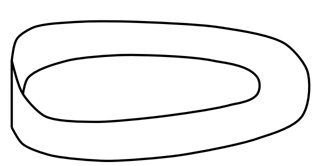
1.21. A compact surface of genus g. Let $f(x) := x(x-1)^2(x-2)^2 \dots (x-(g-1))^2(x-g)$. For small $r > 0$ the set $\{(x, y, z) : (y^2 + f(x))^2 + z^2 = r^2\}$ describes a surface of genus g (topologically a sphere with g handles) in $\mathbb{R}^3$. Visualize this:



1.22. The Moebius strip. It is not the set of zeros of a regular function on an open neighborhood of $\mathbb{R}^n$. Why not? But it may be represented by the following parameterization:

$$f(r, \varphi) := \begin{pmatrix} \cos\varphi(R + r\cos(\varphi/2)) \\ \sin\varphi(R + r\cos(\varphi/2)) \\ r\sin(\varphi/2) \end{pmatrix},$$

$$(r, \varphi) \in (-1, 1) \times [0, 2\pi),$$



where R is quite big.

1.23. Describe an atlas for the real projective plane which consists of three charts (homogeneous coordinates) and compute the chart changings.

Then describe an atlas for the n-dimensional real projective space $P^n(\mathbb{R})$ and compute the chart changes.

1.24. Let $f : L(\mathbb{R}^n, \mathbb{R}^n) \to L(\mathbb{R}^n, \mathbb{R}^n)$ be given by $f(A) := A^\top A$. Where is f of constant rank? What is $f^{-1}(\mathbb{I}_n)$?

1.25. Let $f : L(\mathbb{R}^n, \mathbb{R}^m) \to L(\mathbb{R}^n, \mathbb{R}^n)$, $n < m$, be given by $f(A) := A^\top A$. Where is f of constant rank? What is $f^{-1}(Id_{\mathbb{R}^n})$?

1.26. Let S be a symmetric matrix, i.e., $S(x, y) := x^\top S y$ is a symmetric bilinear form on $\mathbb{R}^n$. Let $f : L(\mathbb{R}^n, \mathbb{R}^n) \to L(\mathbb{R}^n, \mathbb{R}^n)$ be given by $f(A) := A^\top S A$. Where is f of constant rank? What is $f^{-1}(S)$?

1.27. Describe $TS^2 \subset \mathbb{R}^6$.

2. Submersions and Immersions

2.1. Definition. A mapping $f : M \to N$ between manifolds is called a *submersion* at $x \in M$ if the rank of $T_x f : T_x M \to T_{f(x)} N$ equals $\dim N$. Since the rank cannot fall locally (the determinant of a submatrix of the Jacobi matrix is not 0), f is then a submersion in a whole neighborhood of x. The mapping f is said to be a *submersion* if it is a submersion at each $x \in M$.

2.2. Lemma. *If $f : M \to N$ is a submersion at $x \in M$, then for any chart (V, v) centered at $f(x)$ on N there is chart (U, u) centered at x on M such that $v \circ f \circ u^{-1}$ looks as follows:*

$$(y^1, \dots, y^n, y^{n+1}, \dots, y^m) \mapsto (y^1, \dots, y^n).$$

Proof. Use the inverse function theorem once: Apply the argument from the beginning of (1.13) to $v \circ f \circ u_1^{-1}$ for some chart (U_1, u_1) centered at the point x. □

2.3. Corollary. *Any submersion $f : M \to N$ is open: For each open $U \subset M$ the set $f(U)$ is open in N.* □

2.4. Definition. A triple (M, p, N), where $p : M \to N$ is a surjective submersion, is called a *fibered manifold*. The manifold M is called the *total space* and N is called the *base*.

A fibered manifold admits local sections: For each $x \in M$ there is an open neighborhood U of $p(x)$ in N and a smooth mapping $s : U \to M$ with $p \circ s = Id_U$ and $s(p(x)) = x$.

The existence of local sections in turn implies the following universal property:

$$\begin{array}{ccc} M & & \\ {\scriptstyle p}\big\downarrow & \searrow & \\ N & \xrightarrow{\ f\ } & P. \end{array}$$

If (M, p, N) is a fibered manifold and $f : N \to P$ is a mapping into some further manifold such that $f \circ p : M \to P$ is smooth, then f is smooth.

2.5. Definition. A smooth mapping $f : M \to N$ is called an *immersion at* $x \in M$ if the rank of $T_x f : T_x M \to T_{f(x)} N$ equals $\dim M$. Since the rank is maximal at x and cannot fall locally, f is an immersion on a whole neighborhood of x. The mapping f is called an immersion if it is so at every $x \in M$.

2.6. Lemma. *If $f : M \to N$ is an immersion, then for any chart (U, u) centered at $x \in M$ there is a chart (V, v) centered at $f(x)$ on N such that $v \circ f \circ u^{-1}$ has the form*

$$(y^1, \dots, y^m) \mapsto (y^1, \dots, y^m, 0, \dots, 0).$$

Proof. Use the inverse function theorem. □

2.7. Corollary. *If $f : M \to N$ is an immersion, then for any $x \in M$ there is an open neighborhood U of $x \in M$ such that $f(U)$ is a submanifold of N and $f|U : U \to f(U)$ is a diffeomorphism.* □

2.8. Corollary. *If an injective immersion $i : M \to N$ is a homeomorphism onto its image, then $i(M)$ is a submanifold of N.*

Proof. Use (2.7). □

2.9. Definition. If $i : M \to N$ is an injective immersion, then (M, i) is called an *immersed submanifold* of N.

A submanifold is an immersed submanifold, but the converse is wrong in general. The structure of an immersed submanifold (M, i) is in general not determined by the subset $i(M) \subset N$. All this is illustrated by the following example. Consider the curve $\gamma(t) = (\sin^3 t, \sin t. \cos t)$ in $\mathbb{R}^2$. Then $((-\pi, \pi), \gamma|(-\pi, \pi))$ and $((0, 2\pi), \gamma|(0, 2\pi))$ are two different immersed submanifolds, but the image of the embedding is in both cases just the figure eight.

2.10. Let M be a submanifold of N. Then the embedding $i : M \to N$ is an injective immersion with the following property:

(1) *For any manifold Z a mapping $f : Z \to M$ is smooth if and only if $i \circ f : Z \to N$ is smooth.*

There are injective immersions without property (1); see (2.9).

We want to determine all injective immersions $i : M \to N$ with property (1). To require that i is a homeomorphism onto its image is too strong as (2.11) below shows. To look for all smooth mappings $i : M \to N$ with property (2.10.1) (initial mappings in categorical terms) is too difficult as remark (2.12) below shows.

2.11. Example. We consider the 2-dimensional torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. Then the quotient mapping $\pi : \mathbb{R}^2 \to \mathbb{T}^2$ is a covering map, so locally a diffeomorphism. Let us also consider the mapping $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t, \alpha.t)$, where α is irrational. Then $\pi \circ f : \mathbb{R} \to \mathbb{T}^2$ is an injective immersion with dense image, and it is obviously not a homeomorphism onto its image. But $\pi \circ f$ has property (2.10.1), which follows from the fact that π is a covering map.

2.12. Remark. If $f : \mathbb{R} \to \mathbb{R}$ is a function such that f^p and f^q are smooth for some p, q which are relatively prime in $\mathbb{N}$, then f itself turns out to be smooth; see [**97**]. So the mapping $i : t \mapsto \binom{t^p}{t^q}$, $\mathbb{R} \to \mathbb{R}^2$, has property (2.10.1), but i is not an immersion at 0.

In [**98**] all germs of mappings at 0 with property (2.10.1) are characterized as in the following way: Let $g : (\mathbb{R}, 0) \to (\mathbb{R}^n, 0)$ be a germ of a C^∞-curve, $g(t) = (g_1(t), \dots, g_n(t))$. Without loss we may suppose that g is not infinitely flat at 0, so that $g_1(t) = t^r$ for $r \in \mathbb{N}$ after a suitable change of coordinates. Then *g has property* (2.10.1) *near 0 if and only if the Taylor series of g is not contained in any $\mathbb{R}^n[[t^s]]$ for $s \geq 2$.*

2.13. Definition. For an arbitrary subset A of a manifold N and $x_0 \in A$ let $C_{x_0}(A)$ denote the set of all $x \in A$ which can be joined to x_0 by a smooth curve in M lying in A.

A subset M in a manifold N is called an *initial submanifold* of dimension m if the following property is true:

(1) *For each $x \in M$ there exists a chart (U, u) centered at x on N such that $u(C_x(U \cap M)) = u(U) \cap (\mathbb{R}^m \times 0)$.*

The following three lemmas explain the name initial submanifold.

2.14. Lemma. *Let $f : M \to N$ be an injective immersion between manifolds with the universal property* (2.10.1)*. Then $f(M)$ is an initial submanifold of N.*

Proof. Let $x \in M$. By (2.6) we may choose a chart (V, v) centered at $f(x)$ on N and another chart (W, w) centered at x on M such that

$$(v \circ f \circ w^{-1})(y^1, \dots, y^m) = (y^1, \dots, y^m, 0, \dots, 0).$$

Let $r > 0$ be small enough such that $\{y \in \mathbb{R}^m : |y| < 2r\} \subset w(W)$ and also $\{z \in \mathbb{R}^n : |z| < 2r\} \subset v(V)$. Put

$$U := v^{-1}(\{z \in \mathbb{R}^n : |z| < r\}) \subset N,$$
$$W_1 := w^{-1}(\{y \in \mathbb{R}^m : |y| < r\}) \subset M.$$

We claim that $(U, u = v|U)$ satisfies the condition of (2.13.1).

$$\begin{aligned} u^{-1}(u(U) \cap (\mathbb{R}^m \times 0)) &= u^{-1}(\{(y^1, \dots, y^m, 0 \dots, 0) : |y| < r\}) \\ &= f \circ w^{-1} \circ (u \circ f \circ w^{-1})^{-1}(\{(y^1, \dots, y^m, 0 \dots, 0) : |y| < r\}) \\ &= f \circ w^{-1}(\{y \in \mathbb{R}^m : |y| < r\}) = f(W_1) \subseteq C_{f(x)}(U \cap f(M)), \end{aligned}$$

since $f(W_1) \subseteq U \cap f(M)$ and $f(W_1)$ is C^∞-contractible.

Now let conversely $z \in C_{f(x)}(U \cap f(M))$. By definition there is a smooth curve $c : [0,1] \to N$ with $c(0) = f(x)$, $c(1) = z$, and $c([0,1]) \subseteq U \cap f(M)$. By property (2.10.1) the unique curve $\bar{c} : [0,1] \to M$ with $f \circ \bar{c} = c$ is smooth.

We claim that $\bar{c}([0,1]) \subseteq W_1$. If not, then there is some $t \in [0,1]$ with $\bar{c}(t) \in w^{-1}(\{y \in \mathbb{R}^m : r \le |y| < 2r\})$ since $\bar{c}$ is smooth and thus continuous. But then we have

$$\begin{aligned} (v \circ f)(\bar{c}(t)) &\in (v \circ f \circ w^{-1})(\{y \in \mathbb{R}^m : r \le |y| < 2r\}) \\ &= \{(y, 0) \in \mathbb{R}^m \times 0 : r \le |y| < 2r\} \subseteq \{z \in \mathbb{R}^n : r \le |z| < 2r\}. \end{aligned}$$

This means $(v \circ f \circ \bar{c})(t) = (v \circ c)(t) \in \{z \in \mathbb{R}^n : r \le |z| < 2r\}$, so $c(t) \notin U$, a contradiction.

So $\bar{c}([0,1]) \subseteq W_1$; thus $\bar{c}(1) = f^{-1}(z) \in W_1$ and $z \in f(W_1)$. Consequently we have $C_{f(x)}(U \cap f(M)) = f(W_1)$ and finally $f(W_1) = u^{-1}(u(U) \cap (\mathbb{R}^m \times 0))$ by the first part of the proof. □

2.15. Lemma. *Let M be an initial submanifold of a manifold N. Then there is a unique C^∞-manifold structure on M such that the injection $i : M \to N$ is an injective immersion with property* (2.10.1)*:*

(1) *For any manifold Z a mapping $f : Z \to M$ is smooth if and only if $i \circ f : Z \to N$ is smooth.*

The connected components of M are separable (but there may be uncountably many of them).

Proof. We use the sets $C_x(U_x \cap M)$ as charts for M, where $x \in M$ and (U_x, u_x) is a chart for N centered at x with the property required in (2.13.1). Then the chart changings are smooth since they are just restrictions of the chart changings on N. But the sets $C_x(U_x \cap M)$ are not open in the induced topology on M in general. So the identification topology with respect to the charts $(C_x(U_x \cap M), u_x)_{x\in M}$ yields a topology on M which is finer than the induced topology, so it is Hausdorff. Clearly $i : M \to N$ is then an injective immersion. Uniqueness of the smooth structure follows from the universal property (1) which we prove now: For $z \in Z$ we choose a chart (U, u) on N, centered at $f(z)$, such that $u(C_{f(z)}(U \cap M)) = u(U) \cap (\mathbb{R}^m \times 0)$. Then $f^{-1}(U)$ is open in Z and contains a chart (V, v) centered at z on Z with $v(V)$ a ball. Then $f(V)$ is C^∞-contractible in $U \cap M$, so $f(V) \subseteq C_{f(z)}(U \cap M)$, and $(u|C_{f(z)}(U \cap M)) \circ f \circ v^{-1} = u \circ f \circ v^{-1}$ is smooth.

Finally note that N admits a Riemann metric (22.1) which induces one on M, so each connected component of M is separable, by (1.1.4). □

2.16. Transversal mappings. Let M_1, M_2, and N be manifolds and let $f_i : M_i \to N$ be smooth mappings for $i = 1, 2$. We say that f_1 and f_2 are *transversal* at $y \in N$ if

$$\operatorname{im} T_{x_1} f_1 + \operatorname{im} T_{x_2} f_2 = T_y N \quad \text{whenever} \quad f_1(x_1) = f_2(x_2) = y.$$

Note that they are transversal at any y which is not in $f_1(M_1)$ or not in $f_2(M_2)$. The mappings f_1 and f_2 are simply said to be *transversal* if they are transversal at every $y \in N$.

If P is an initial submanifold of N with embedding $i : P \to N$, then a mapping $f : M \to N$ is said to be transversal to P if i and f are transversal.

Lemma. *In this case $f^{-1}(P)$ is an initial submanifold of M with the same codimension in M as P has in N; or $f^{-1}(P)$ is the empty set. If P is a submanifold, then also $f^{-1}(P)$ is a submanifold.*

Proof. Let $x \in f^{-1}(P)$ and let (U, u) be an initial submanifold chart for P centered at $f(x)$ on N, i.e., $u(C_{f(x)}(U \cap P)) = u(U) \cap (\mathbb{R}^p \times 0)$. Then the mapping

$$M \supseteq f^{-1}(U) \xrightarrow{f} U \xrightarrow{u} u(U) \subseteq \mathbb{R}^p \times \mathbb{R}^{n-p} \xrightarrow{\mathrm{pr}_2} \mathbb{R}^{n-p}$$

is a submersion at x since f is transversal to P. So by lemma (2.2) there is a chart (V, v) on M centered at x such that we have

$$(\mathrm{pr}_2 \circ u \circ f \circ v^{-1})(y^1, \dots, y^{n-p}, \dots, y^m) = (y^1, \dots, y^{n-p}).$$

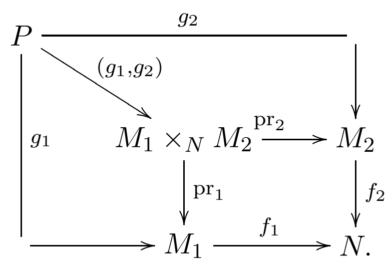
But then $z \in C_x(f^{-1}(P) \cap V)$ if and only if $v(z) \in v(V) \cap (0 \times \mathbb{R}^{m-n+p})$, so $v(C_x(f^{-1}(P) \cap V)) = v(V) \cap (0 \times \mathbb{R}^{m-n+p})$. □

2.17. Corollary. *If $f_1 : M_1 \to N$ and $f_2 : M_2 \to N$ are smooth and transversal, then the topological pullback*

$$M_1 \underset{(f_1,N,f_2)}{\times} M_2 = M_1 \times_N M_2 := \{(x_1, x_2) \in M_1 \times M_2 : f_1(x_1) = f_2(x_2)\}$$

is a submanifold of $M_1 \times M_2$, and it has the following universal property:

> *For any smooth mappings $g_1 : P \to M_1$ and $g_2 : P \to M_2$ with $f_1 \circ g_1 = f_2 \circ g_2$ there is a unique smooth mapping $(g_1, g_2) : P \to M_1 \times_N M_2$ with $\mathrm{pr}_1 \circ (g_1, g_2) = g_1$ and $\mathrm{pr}_2 \circ (g_1, g_2) = g_2$.*



This is also called the pullback property in the category $\mathcal{M}f$ of smooth manifolds and smooth mappings. So one may say that transversal pullbacks exist in the category $\mathcal{M}f$. But there also exist pullbacks which are not transversal.

Proof. $M_1 \times_N M_2 = (f_1 \times f_2)^{-1}(\Delta)$, where $f_1 \times f_2 : M_1 \times M_2 \to N \times N$ and where Δ is the diagonal of $N \times N$, and $f_1 \times f_2$ is transversal to Δ if and only if f_1 and f_2 are transversal. □

3. Vector Fields and Flows

3.1. Definition. A *vector field* X on a manifold M is a smooth section of the tangent bundle; so $X : M \to TM$ is smooth and $\pi_M \circ X = Id_M$. A *local vector field* is a smooth section which is defined on an open subset only. We denote the set of all vector fields by $\mathfrak{X}(M)$. With pointwise addition and scalar multiplication $\mathfrak{X}(M)$ becomes a vector space.

Example. Let (U, u) be a chart on M. Then the $\frac{\partial}{\partial u^i} : U \to TM|U$, $x \mapsto \frac{\partial}{\partial u^i}|_x$, described in (1.8), are local vector fields defined on U.

Lemma. *If X is a vector field on M and (U, u) is a chart on M and $x \in U$, then we have $X(x) = \sum_{i=1}^m X(x)(u^i)\frac{\partial}{\partial u^i}|_x$. We write $X|U = \sum_{i=1}^m X(u^i)\frac{\partial}{\partial u^i}$.* □

3.2. The vector fields $(\frac{\partial}{\partial u^i})_{i=1}^m$ on U, where (U,u) is a chart on M, form a *holonomic frame field*. By a *frame field* on some open set $V \subset M$ we mean $m = \dim M$ vector fields $s_i \in \mathfrak{X}(U)$ such that $s_1(x), \dots, s_m(x)$ is a linear basis of T_xM for each $x \in V$. A frame field is said to be *holonomic* if $s_i = \frac{\partial}{\partial v^i}$ for some chart (V,v). If no such chart may be found locally, the frame field is called *anholonomic*.

With the help of partitions of unity and holonomic frame fields one may construct 'many' vector fields on M. In particular the values of a vector field can be arbitrarily preassigned on a discrete set $\{x_i\} \subset M$.

3.3. Lemma. *The space $\mathfrak{X}(M)$ of vector fields on M coincides canonically with the space of all derivations of the algebra $C^\infty(M)$ of smooth functions, i.e., those $\mathbb{R}$-linear operators $D : C^\infty(M) \to C^\infty(M)$ with*

$$D(fg) = D(f)g + fD(g).$$

Proof. Clearly each vector field $X \in \mathfrak{X}(M)$ defines a derivation (again called X; later sometimes called $\mathcal{L}_X$) of the algebra $C^\infty(M)$ by stipulating $X(f)(x) := X(x)(f) = df(X(x))$.

If conversely a derivation D of $C^\infty(M)$ is given, for any $x \in M$ we consider $D_x : C^\infty(M) \to \mathbb{R}$, $D_x(f) = D(f)(x)$. Then D_x is a derivation at x of $C^\infty(M)$ in the sense of (1.7), so $D_x = X_x$ for some $X_x \in T_xM$. In this way we get a section $X : M \to TM$. If (U,u) is a chart on M, we have $D_x = \sum_{i=1}^m X(x)(u^i)\frac{\partial}{\partial u^i}|_x$ by (1.7). Choose V open in M, $V \subset \overline{V} \subset U$, and $\varphi \in C^\infty(M,\mathbb{R})$ such that $\operatorname{supp}(\varphi) \subset U$ and $\varphi|V = 1$. Then $\varphi \cdot u^i \in C^\infty(M)$ and $(\varphi u^i)|V = u^i|V$. So $D(\varphi u^i)(x) = X(x)(\varphi u^i) = X(x)(u^i)$ and $X|V = \sum_{i=1}^m D(\varphi u^i)|V \cdot \frac{\partial}{\partial u^i}|V$ is smooth. □

3.4. The Lie bracket. By lemma (3.3) we can identify $\mathfrak{X}(M)$ with the vector space of all derivations of the algebra $C^\infty(M)$, which we will do without any notational change in the following.

If X, Y are two vector fields on M, then the mapping $f \mapsto X(Y(f)) - Y(X(f))$ is again a derivation of $C^\infty(M)$, as a simple computation shows. Thus there is a unique vector field $[X,Y] \in \mathfrak{X}(M)$ such that $[X,Y](f) = X(Y(f)) - Y(X(f))$ holds for all $f \in C^\infty(M)$.

In a local chart (U,u) on M one easily checks that for $X|U = \sum X^i\frac{\partial}{\partial u^i}$ and $Y|U = \sum Y^i\frac{\partial}{\partial u^i}$ we have

$$\Big[\sum_i X^i\tfrac{\partial}{\partial u^i}, \sum_j Y^j\tfrac{\partial}{\partial u^j}\Big] = \sum_{i,j}\big(X^i(\tfrac{\partial}{\partial u^i}Y^j) - Y^i(\tfrac{\partial}{\partial u^i}X^j)\big)\tfrac{\partial}{\partial u^j},$$

since second partial derivatives commute. The $\mathbb{R}$-bilinear mapping

$$[\quad,\quad] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$$

is called the *Lie bracket*. Note also that $\mathfrak{X}(M)$ is a module over the algebra $C^\infty(M)$ by pointwise multiplication $(f, X) \mapsto fX$.

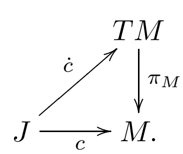
Theorem. *The Lie bracket* $[\ \ ,\ \ ] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ *has the following properties:*

$$[X, Y] = -[Y, X],$$
$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]], \quad \textit{the Jacobi identity,}$$
$$[fX, Y] = f[X, Y] - (Yf)X,$$
$$[X, fY] = f[X, Y] + (Xf)Y.$$

The form of the Jacobi identity we have chosen says that $\operatorname{ad}(X) = [X,\ \ ]$ is a derivation for the Lie algebra $(\mathfrak{X}(M), [\ \ ,\ \ ])$. The pair $(\mathfrak{X}(M), [\ \ ,\ \ ])$ is the prototype of a *Lie algebra*. The concept of a Lie algebra is one of the most important notions of modern mathematics.

Proof. All these properties are checked easily for the commutator $[X, Y] = X \circ Y - Y \circ X$ in the space of derivations of the algebra $C^\infty(M)$. □

3.5. Integral curves. Let $c : J \to M$ be a smooth curve in a manifold M defined on an interval J. We will use the following notations: $c'(t) = \dot{c}(t) = \frac{d}{dt}c(t) := T_t c.1$. Clearly $c' : J \to TM$ is smooth. We call c' a vector field along c since we have $\pi_M \circ c' = c$:



A smooth curve $c : J \to M$ will be called an *integral curve* or *flow line* of a vector field $X \in \mathfrak{X}(M)$ if $c'(t) = X(c(t))$ holds for all $t \in J$.

3.6. Lemma. *Let X be a vector field on M. Then for any $x \in M$ there is an open interval J_x containing 0 and an integral curve $c_x : J_x \to M$ for X (i.e., $c_x' = X \circ c_x$) with $c_x(0) = x$. If J_x is maximal, then c_x is unique.*

Proof. In a chart (U, u) on M with $x \in U$ the equation $c'(t) = X(c(t))$ is a system ordinary differential equations with initial condition $c(0) = x$. Since X is smooth, there is a unique local solution which even depends smoothly on the initial values, by the theorem of Picard-Lindelöf, [**41**, 10.7.4]. So on M there are always local integral curves. If $J_x = (a, b)$ and $\lim_{t \to b-} c_x(t) =: c_x(b)$ exists in M, there is a unique local solution c_1 defined in an open interval containing b with $c_1(b) = c_x(b)$. By uniqueness of the solution on the intersection of the two intervals, c_1 prolongs c_x to a larger interval. This may be repeated (also on the left hand side of J_x) as long as the limit exists. So if we suppose J_x to be maximal, J_x either equals $\mathbb{R}$ or the integral curve leaves the manifold in finite (parameter-)time in the past or future or both. □

3.7. The flow of a vector field. Let $X \in \mathfrak{X}(M)$ be a vector field. Let us write $\mathrm{Fl}^X_t(x) = \mathrm{Fl}^X(t,x) := c_x(t)$, where $c_x : J_x \to M$ is the maximally defined integral curve of X with $c_x(0) = x$, constructed in lemma (3.6).

Theorem. *For each vector field X on M, the mapping $\mathrm{Fl}^X : \mathcal{D}(X) \to M$ is smooth, where $\mathcal{D}(X) = \bigcup_{x\in M} J_x \times \{x\}$ is an open neighborhood of $0 \times M$ in $\mathbb{R} \times M$. We have*

$$\mathrm{Fl}^X(t+s,x) = \mathrm{Fl}^X(t, \mathrm{Fl}^X(s,x))$$

in the following sense. If the right hand side exists, then the left hand side exists and we have equality. If both t, $s \geq 0$ or both are ≤ 0, and if the left hand side exists, then also the right hand side exists and we have equality.

Proof. As mentioned in the proof of (3.6), $\mathrm{Fl}^X(t,x)$ is smooth in (t,x) for small t, and if it is defined for (t,x), then it is also defined for (s,y) nearby. These are local properties which follow from the theory of ordinary differential equations.

Now let us treat the equation $\mathrm{Fl}^X(t+s,x) = \mathrm{Fl}^X(t, \mathrm{Fl}^X(s,x))$. If the right hand side exists, then we consider the equation

$$\begin{cases} \frac{d}{dt}\mathrm{Fl}^X(t+s,x) = \frac{d}{du}\mathrm{Fl}^X(u,x)|_{u=t+s} = X(\mathrm{Fl}^X(t+s,x)), \\ \mathrm{Fl}^X(t+s,x)|_{t=0} = \mathrm{Fl}^X(s,x). \end{cases}$$

But the unique solution of this is $\mathrm{Fl}^X(t, \mathrm{Fl}^X(s,x))$. So the left hand side exists and equals the right hand side.

If the left hand side exists, let us suppose that $t, s \geq 0$. We put

$$c_x(u) = \begin{cases} 2\,\mathrm{Fl}^X(u,x) & \text{if } u \leq s, \\ \quad\mathrm{Fl}^X(u-s, \mathrm{Fl}^X(s,x)) & \text{if } u \geq s. \end{cases}$$

Then we have

$$\begin{aligned} \tfrac{d}{du}c_x(u) &= \begin{cases} \frac{d}{du}\mathrm{Fl}^X(u,x) = X(\mathrm{Fl}^X(u,x)) \quad \text{for } u \leq s, \\ \frac{d}{du}\mathrm{Fl}^X(u-s,\mathrm{Fl}^X(s,x)) = X(\mathrm{Fl}^X(u-s,\mathrm{Fl}^X(s,x))) \end{cases} \\ &= X(c_x(u)) \quad \text{for } 0 \leq u \leq t+s. \end{aligned}$$

Also $c_x(0) = x$ and on the overlap both definitions coincide by the first part of the proof; thus we conclude that $c_x(u) = \mathrm{Fl}^X(u,x)$ for $0 \leq u \leq t+s$ and we have $\mathrm{Fl}^X(t, \mathrm{Fl}^X(s,x)) = c_x(t+s) = \mathrm{Fl}^X(t+s,x)$.

Now we show that $\mathcal{D}(X)$ is open and Fl^X is smooth on $\mathcal{D}(X)$. We know already that $\mathcal{D}(X)$ is a neighborhood of $0 \times M$ in $\mathbb{R} \times M$ and that Fl^X is smooth near $0 \times M$.

For $x \in M$ let J'_x be the set of all $t \in \mathbb{R}$ such that Fl^X is defined and smooth on an open neighborhood of $[0,t] \times \{x\}$ (respectively on $[t,0] \times \{x\}$ for $t < 0$) in $\mathbb{R} \times M$. We claim that $J'_x = J_x$, which finishes the proof. It suffices to show that J'_x is not empty, open and closed in J_x. It is open by construction, and not empty, since $0 \in J'_x$. If J'_x is not closed in J_x, let $t_0 \in J_x \cap (\overline{J'_x} \setminus J'_x)$ and suppose that $t_0 > 0$, say. By the local existence and smoothness Fl^X exists and is smooth near $[-\varepsilon, \varepsilon] \times \{y := \mathrm{Fl}^X(t_0, x)\}$ in $\mathbb{R} \times M$ for some $\varepsilon > 0$, and by construction Fl^X exists and is smooth near $[0, t_0 - \varepsilon] \times \{x\}$. Since $\mathrm{Fl}^X(-\varepsilon, y) = \mathrm{Fl}^X(t_0 - \varepsilon, x)$, we conclude for t near $[0, t_0 - \varepsilon]$, x' near x, and t' near $[-\varepsilon, \varepsilon]$ that $\mathrm{Fl}^X(t + t', x') = \mathrm{Fl}^X(t', \mathrm{Fl}^X(t, x'))$ exists and is smooth. So $t_0 \in J'_x$, a contradiction. □

3.8. Let $X \in \mathfrak{X}(M)$ be a vector field. Its flow Fl^X is called *global* or *complete* if its domain of definition $\mathcal{D}(X)$ equals $\mathbb{R} \times M$. Then the vector field X itself will be called a *complete vector field*. In this case Fl^X_t is also sometimes called $\exp tX$; it is a diffeomorphism of M. The *support* $\mathrm{supp}(X)$ of a vector field X is the closure of the set $\{x \in M : X(x) \neq 0\}$.

Lemma. *A vector field with compact support on M is complete.*

Proof. Let $K = \mathrm{supp}(X)$ be compact. Then the compact set $0 \times K$ has positive distance to the disjoint closed set $(\mathbb{R} \times M) \setminus \mathcal{D}(X)$ (if it is not empty), so $[-\varepsilon, \varepsilon] \times K \subset \mathcal{D}(X)$ for some $\varepsilon > 0$. If $x \notin K$, then $X(x) = 0$, so $\mathrm{Fl}^X(t, x) = x$ for all t and $\mathbb{R} \times \{x\} \subset \mathcal{D}(X)$. So we have $[-\varepsilon, \varepsilon] \times M \subset \mathcal{D}(X)$. Since $\mathrm{Fl}^X(t + \varepsilon, x) = \mathrm{Fl}^X(t, \mathrm{Fl}^X(\varepsilon, x))$ exists for $|t| \leq \varepsilon$ by theorem (3.7), we have $[-2\varepsilon, 2\varepsilon] \times M \subset \mathcal{D}(X)$ and by repeating this argument we get $\mathbb{R} \times M = \mathcal{D}(X)$. □

So on a compact manifold M each vector field is complete. If M is not compact and of dimension ≥ 2, then in general the set of complete vector fields on M is neither a vector space nor is it closed under the Lie bracket, as the following example on $\mathbb{R}^2$ shows: $X = y\frac{\partial}{\partial x}$ and $Y = \frac{x^2}{2}\frac{\partial}{\partial y}$ are complete, but neither $X + Y$ nor $[X, Y]$ is complete. In general one may embed $\mathbb{R}^2$ as a closed submanifold into M and extend the vector fields X and Y.

3.9. f-related vector fields. If $f : M \to M$ is a diffeomorphism, then for any vector field $X \in \mathfrak{X}(M)$ the mapping $Tf^{-1} \circ X \circ f$ is also a vector field, which we will denote by f^*X. We also put $f_*X := Tf \circ X \circ f^{-1} = (f^{-1})^*X$.

But if $f : M \to N$ is a smooth mapping and $Y \in \mathfrak{X}(N)$ is a vector field, there may or may not exist a vector field $X \in \mathfrak{X}(M)$ such that the following diagram commutes:

(1)
$$\begin{array}{ccc} TM & \xrightarrow{Tf} & TN \\ {\scriptstyle X}\uparrow & & \uparrow{\scriptstyle Y} \\ M & \xrightarrow{f} & N. \end{array}$$

Definition. Let $f : M \to N$ be a smooth mapping. Two vector fields $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ are called *f-related* if $Tf \circ X = Y \circ f$ holds, i.e., if diagram (1) commutes.

Example. If $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ and if $X \times Y \in \mathfrak{X}(M \times N)$ is given by $(X \times Y)(x, y) = (X(x), Y(y))$, then we have:

(2) $X \times Y$ and X are pr_1-related.

(3) $X \times Y$ and Y are pr_2-related.

(4) X and $X \times Y$ are $\mathrm{ins}(y)$-related if and only if $Y(y) = 0$, where the mapping $\mathrm{ins}(y) : M \to M \times N$ is given by $\mathrm{ins}(y)(x) = (x, y)$.

3.10. Lemma. *Consider vector fields $X_i \in \mathfrak{X}(M)$ and $Y_i \in \mathfrak{X}(N)$ for $i = 1, 2$, and a smooth mapping $f : M \to N$. If X_i and Y_i are f-related for $i = 1, 2$, then also $\lambda_1 X_1 + \lambda_2 X_2$ and $\lambda_1 Y_1 + \lambda_2 Y_2$ are f-related, and also $[X_1, X_2]$ and $[Y_1, Y_2]$ are f-related.*

Proof. The first assertion is immediate. To prove the second, we choose $h \in C^\infty(N)$. Then by assumption we have $Tf \circ X_i = Y_i \circ f$; thus:

$$\begin{aligned}(X_i(h \circ f))(x) &= X_i(x)(h \circ f) = (T_x f.X_i(x))(h) \\ &= (Tf \circ X_i)(x)(h) = (Y_i \circ f)(x)(h) = Y_i(f(x))(h) = (Y_i(h))(f(x)),\end{aligned}$$

so $X_i(h \circ f) = (Y_i(h)) \circ f$, and we may continue:

$$\begin{aligned}[X_1, X_2](h \circ f) &= X_1(X_2(h \circ f)) - X_2(X_1(h \circ f)) \\ &= X_1(Y_2(h) \circ f) - X_2(Y_1(h) \circ f) \\ &= Y_1(Y_2(h)) \circ f - Y_2(Y_1(h)) \circ f = [Y_1, Y_2](h) \circ f.\end{aligned}$$

But this means $Tf \circ [X_1, X_2] = [Y_1, Y_2] \circ f$. □

3.11. Corollary. *If $f : M \to N$ is a local diffeomorphism (so $(T_x f)^{-1}$ makes sense for each $x \in M$), then for $Y \in \mathfrak{X}(N)$ a vector field $f^*Y \in \mathfrak{X}(M)$ is defined by $(f^*Y)(x) = (T_x f)^{-1}.Y(f(x))$. The linear mapping $f^* : \mathfrak{X}(N) \to \mathfrak{X}(M)$ is then a Lie algebra homomorphism, i.e.,*

$$f^*[Y_1, Y_2] = [f^*Y_1, f^*Y_2].$$

3.12. The Lie derivative of functions. For a vector field $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$ we define $\mathcal{L}_X f \in C^\infty(M)$ by

$$\mathcal{L}_X f(x) := \tfrac{d}{dt}|_0 f(\operatorname{Fl}^X(t,x)) \quad \text{or}$$
$$\mathcal{L}_X f := \tfrac{d}{dt}|_0 (\operatorname{Fl}^X_t)^* f = \tfrac{d}{dt}|_0 (f \circ \operatorname{Fl}^X_t).$$

Since $\operatorname{Fl}^X(t,x)$ is defined for small t, for any $x \in M$, the expressions above make sense.

Lemma. *We have*

$$\tfrac{d}{dt}(\operatorname{Fl}^X_t)^* f = (\operatorname{Fl}^X_t)^* X(f) = X((\operatorname{Fl}^X_t)^* f);$$

in particular for $t = 0$ we have $\mathcal{L}_X f = X(f) = df(X)$.

Proof. We have

$$\tfrac{d}{dt}(\operatorname{Fl}^X_t)^* f(x) = df(\tfrac{d}{dt}\operatorname{Fl}^X(t,x)) = df(X(\operatorname{Fl}^X(t,x))) = (\operatorname{Fl}^X_t)^*(Xf)(x).$$

From this we get $\mathcal{L}_X f = X(f) = df(X)$ and then in turn

$$\tfrac{d}{dt}(\operatorname{Fl}^X_t)^* f = \tfrac{d}{ds}|_0 (\operatorname{Fl}^X_t \circ \operatorname{Fl}^X_s)^* f = \tfrac{d}{ds}|_0 (\operatorname{Fl}^X_s)^* (\operatorname{Fl}^X_t)^* f = X((\operatorname{Fl}^X_t)^* f). \quad \square$$

3.13. The Lie derivative for vector fields. For $X, Y \in \mathfrak{X}(M)$ we define $\mathcal{L}_X Y \in \mathfrak{X}(M)$ by

$$\mathcal{L}_X Y := \tfrac{d}{dt}|_0 (\operatorname{Fl}^X_t)^* Y = \tfrac{d}{dt}|_0 (T(\operatorname{Fl}^X_{-t}) \circ Y \circ \operatorname{Fl}^X_t),$$

and call it the *Lie derivative* of Y along X.

Lemma. *We have*

$$\mathcal{L}_X Y = [X, Y],$$
$$\tfrac{d}{dt}(\operatorname{Fl}^X_t)^* Y = (\operatorname{Fl}^X_t)^* \mathcal{L}_X Y = (\operatorname{Fl}^X_t)^*[X, Y] = \mathcal{L}_X (\operatorname{Fl}^X_t)^* Y = [X, (\operatorname{Fl}^X_t)^* Y].$$

Proof. For $f \in C^\infty(M)$ consider the mapping $\alpha(t,s) := Y(\operatorname{Fl}^X(t,x))(f \circ \operatorname{Fl}^X_s)$, which is locally defined near 0. It satisfies

$$\begin{aligned}
\alpha(t,0) &= Y(\operatorname{Fl}^X(t,x))(f),\\
\alpha(0,s) &= Y(x)(f \circ \operatorname{Fl}^X_s),\\
\tfrac{\partial}{\partial t}\alpha(0,0) &= \partial|_0 Y(\operatorname{Fl}^X(t,x))(f) = \partial|_0 (Yf)(\operatorname{Fl}^X(t,x)) = X(x)(Yf),\\
\tfrac{\partial}{\partial s}\alpha(0,0) &= \tfrac{\partial}{\partial s}|_0 Y(x)(f \circ \operatorname{Fl}^X_s) = Y(x)\tfrac{\partial}{\partial s}|_0 (f \circ \operatorname{Fl}^X_s) = Y(x)(Xf).
\end{aligned}$$

But on the other hand we have

$$\begin{aligned}
\tfrac{\partial}{\partial u}|_0 \alpha(u, -u) &= \tfrac{\partial}{\partial u}|_0 Y(\operatorname{Fl}^X(u,x))(f \circ \operatorname{Fl}^X_{-u})\\
&= \tfrac{\partial}{\partial u}|_0 \left(T(\operatorname{Fl}^X_{-u}) \circ Y \circ \operatorname{Fl}^X_u\right)_x (f) = (\mathcal{L}_X Y)_x(f),
\end{aligned}$$

so the first assertion follows. For the second claim we compute as follows:

$$\begin{aligned}\tfrac{\partial}{\partial t}(\mathrm{Fl}^X_t)^*Y &= \tfrac{\partial}{\partial s}|_0\left(T(\mathrm{Fl}^X_{-t})\circ T(\mathrm{Fl}^X_{-s})\circ Y\circ \mathrm{Fl}^X_s\circ \mathrm{Fl}^X_t\right)\\ &= T(\mathrm{Fl}^X_{-t})\circ \tfrac{\partial}{\partial s}|_0\left(T(\mathrm{Fl}^X_{-s})\circ Y\circ \mathrm{Fl}^X_s\right)\circ \mathrm{Fl}^X_t\\ &= T(\mathrm{Fl}^X_{-t})\circ [X,Y]\circ \mathrm{Fl}^X_t = (\mathrm{Fl}^X_t)^*[X,Y].\\ \tfrac{\partial}{\partial t}(\mathrm{Fl}^X_t)^*Y &= \tfrac{\partial}{\partial s}|_0(\mathrm{Fl}^X_s)^*(\mathrm{Fl}^X_t)^*Y = \mathcal{L}_X(\mathrm{Fl}^X_t)^*Y. \quad \square\end{aligned}$$

3.14. Lemma. *Let $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ be f-related vector fields for a smooth mapping $f : M \to N$. Then we have $f \circ \mathrm{Fl}^X_t = \mathrm{Fl}^Y_t \circ f$, whenever both sides are defined. In particular, if f is a diffeomorphism, we have $\mathrm{Fl}^{f^*Y}_t = f^{-1}\circ \mathrm{Fl}^Y_t \circ f$.*

Proof. We have $\frac{d}{dt}(f\circ \mathrm{Fl}^X_t) = Tf\circ \frac{d}{dt}\mathrm{Fl}^X_t = Tf\circ X\circ \mathrm{Fl}^X_t = Y\circ f\circ Fl^X_t$ and $f(\mathrm{Fl}^X(0,x)) = f(x)$. So $t\mapsto f(\mathrm{Fl}^X(t,x))$ is an integral curve of the vector field Y on N with initial value $f(x)$, so we have $f(\mathrm{Fl}^X(t,x)) = \mathrm{Fl}^Y(t,f(x))$ or $f\circ \mathrm{Fl}^X_t = \mathrm{Fl}^Y_t\circ f$. $\square$

3.15. Corollary. *Let $X, Y \in \mathfrak{X}(M)$. Then the following assertions are equivalent:*

(1) $\mathcal{L}_X Y = [X,Y] = 0$.
(2) $(\mathrm{Fl}^X_t)^*Y = Y$, *wherever defined.*
(3) $\mathrm{Fl}^X_t\circ \mathrm{Fl}^Y_s = \mathrm{Fl}^Y_s\circ \mathrm{Fl}^X_t$, *wherever defined.*

Proof. (1) $\Leftrightarrow$ (2) is immediate from lemma (3.13). To see (2) $\Leftrightarrow$ (3), we note that $\mathrm{Fl}^X_t\circ \mathrm{Fl}^Y_s = \mathrm{Fl}^Y_s\circ \mathrm{Fl}^X_t$ if and only if $\mathrm{Fl}^Y_s = \mathrm{Fl}^X_{-t}\circ \mathrm{Fl}^Y_s\circ \mathrm{Fl}^X_t = \mathrm{Fl}^{(\mathrm{Fl}^X_t)^*Y}_s$ by lemma (3.14); and this in turn is equivalent to $Y = (\mathrm{Fl}^X_t)^*Y$. $\square$

3.16. Theorem. *Let M be a manifold, let $\varphi^i : \mathbb{R}\times M \supset U_{\varphi^i} \to M$ be smooth mappings for $i = 1,\dots,k$ where each U_{φ^i} is an open neighborhood of $\{0\}\times M$ in $\mathbb{R}\times M$, such that each φ^i_t is a diffeomorphism on its domain, $\varphi^i_0 = Id_M$, and $\partial|_0\varphi^i_t = X_i \in \mathfrak{X}(M)$. We put $[\varphi^i,\varphi^j]_t = [\varphi^i_t,\varphi^j_t] := (\varphi^j_t)^{-1}\circ(\varphi^i_t)^{-1}\circ\varphi^j_t\circ\varphi^i_t$. Then for each formal bracket expression P of length k we have*

$$0 = \tfrac{\partial^\ell}{\partial t^\ell}|_0 P(\varphi^1_t,\dots,\varphi^k_t) \quad \text{for } 1\le \ell < k,$$
$$P(X_1,\dots,X_k) = \tfrac{1}{k!}\tfrac{\partial^k}{\partial t^k}|_0 P(\varphi^1_t,\dots,\varphi^k_t) \in \mathfrak{X}(M)$$

in the sense explained in step 2 of the proof. In particular we have for vector fields $X, Y \in \mathfrak{X}(M)$

$$0 = \partial|_0(\mathrm{Fl}^Y_{-t}\circ \mathrm{Fl}^X_{-t}\circ \mathrm{Fl}^Y_t\circ \mathrm{Fl}^X_t),$$
$$[X,Y] = \tfrac{1}{2}\tfrac{\partial^2}{\partial t^2}|_0(\mathrm{Fl}^Y_{-t}\circ \mathrm{Fl}^X_{-t}\circ \mathrm{Fl}^Y_t\circ \mathrm{Fl}^X_t).$$

Proof. Step 1. Let $c : \mathbb{R} \to M$ be a smooth curve. If $c(0) = x \in M$, $c'(0) = 0, \dots, c^{(k-1)}(0) = 0$, then $c^{(k)}(0)$ is a well defined tangent vector in T_xM which is given by the derivation $f \mapsto (f \circ c)^{(k)}(0)$ at x. Namely, we have

$$((f.g) \circ c)^{(k)}(0) = ((f \circ c).(g \circ c))^{(k)}(0) = \sum_{j=0}^{k} \binom{k}{j} (f \circ c)^{(j)}(0)(g \circ c)^{(k-j)}(0)$$
$$= (f \circ c)^{(k)}(0)g(x) + f(x)(g \circ c)^{(k)}(0),$$

since all other summands vanish: $(f \circ c)^{(j)}(0) = 0$ for $1 \le j < k$.

Step 2. Let $\varphi : \mathbb{R} \times M \supset U_\varphi \to M$ be a smooth mapping where U_φ is an open neighborhood of $\{0\} \times M$ in $\mathbb{R} \times M$, such that each φ_t is a diffeomorphism on its domain and $\varphi_0 = Id_M$. We say that φ_t is a *curve of local diffeomorphisms* through Id_M.

From step 1 we see that if $\frac{\partial^j}{\partial t^j}|_0 \varphi_t = 0$ for all $1 \le j < k$, then $X := \frac{1}{k!} \frac{\partial^k}{\partial t^k}|_0 \varphi_t$ is a well defined vector field on M. We say that X is the *first nonvanishing derivative* at 0 of the curve φ_t of local diffeomorphisms. We may paraphrase this as $(\partial_t^k|_0 \varphi_t^*)f = k! \mathcal{L}_X f$.

Claim 3. Let φ_t, ψ_t be curves of local diffeomorphisms through Id_M and let $f \in C^\infty(M)$. Then we have

$$\partial_t^k|_0 (\varphi_t \circ \psi_t)^* f = \partial_t^k|_0 (\psi_t^* \circ \varphi_t^*) f = \sum_{j=0}^{k} \binom{k}{j} (\partial_t^j|_0 \psi_t^*)(\partial_t^{k-j}|_0 \varphi_t^*) f.$$

Also the multinomial version of this formula holds:

$$\partial_t^k|_0 (\varphi_t^1 \circ \dots \circ \varphi_t^\ell)^* f = \sum_{j_1 + \dots + j_\ell = k} \frac{k!}{j_1! \dots j_\ell!} (\partial_t^{j_\ell}|_0 (\varphi_t^\ell)^*) \dots (\partial_t^{j_1}|_0 (\varphi_t^1)^*) f.$$

We only show the binomial version. For a function $h(t,s)$ of two variables we have

$$\partial_t^k h(t,t) = \sum_{j=0}^{k} \binom{k}{j} \partial_t^j \partial_s^{k-j} h(t,s)|_{s=t},$$

since for $h(t,s) = f(t)g(s)$ this is just a consequence of the Leibniz rule, and linear combinations of such decomposable tensors are dense in the space of all functions of two variables in the compact C^∞-topology, so that by continuity the formula holds for all functions. In the following form it implies the claim:

$$\partial_t^k|_0 f(\varphi(t, \psi(t,x))) = \sum_{j=0}^{k} \binom{k}{j} \partial_t^j \partial_s^{k-j} f(\varphi(t, \psi(s,x)))|_{t=s=0}.$$

Claim 4. Let φ_t be a curve of local diffeomorphisms through Id_M with first nonvanishing derivative $k!X = \partial_t^k|_0 \varphi_t$. Then the inverse curve of local diffeomorphisms φ_t^{-1} has first nonvanishing derivative $-k!X = \partial_t^k|_0\varphi_t^{-1}$, for we have $\varphi_t^{-1}\circ\varphi_t = Id$, so by claim 3 we get for $1 \le j \le k$

$$\begin{aligned}0 = \partial_t^j|_0(\varphi_t^{-1}\circ\varphi_t)^*f &= \sum_{i=0}^{j}\tbinom{j}{i}(\partial_t^i|_0\varphi_t^*)(\partial_t^{j-i}(\varphi_t^{-1})^*)f\\ &= \partial_t^j|_0\varphi_t^*(\varphi_0^{-1})^*f + \varphi_0^*\partial_t^j|_0(\varphi_t^{-1})^*f,\end{aligned}$$

i.e., $\partial_t^j|_0\varphi_t^*f = -\partial_t^j|_0(\varphi_t^{-1})^*f$ as required.

Claim 5. Let φ_t be a curve of local diffeomorphisms through Id_M with first nonvanishing derivative $m!X = \partial_t^m|_0\varphi_t$, and let ψ_t be a curve of local diffeomorphisms through Id_M with first nonvanishing derivative $n!Y = \partial_t^n|_0\psi_t$.

Then the curve of local diffeomorphisms $[\varphi_t,\psi_t] = \psi_t^{-1}\circ\varphi_t^{-1}\circ\psi_t\circ\varphi_t$ has first nonvanishing derivative

$$(m+n)![X,Y] = \partial_t^{m+n}|_0[\varphi_t,\psi_t].$$

From this claim the theorem follows.

By the multinomial version of claim 3 we have

$$\begin{aligned}A_Nf :&= \partial_t^N|_0(\psi_t^{-1}\circ\varphi_t^{-1}\circ\psi_t\circ\varphi_t)^*f\\ &= \sum_{i+j+k+\ell=N}\frac{N!}{i!j!k!\ell!}(\partial_t^i|_0\varphi_t^*)(\partial_t^j|_0\psi_t^*)(\partial_t^k|_0(\varphi_t^{-1})^*)(\partial_t^\ell|_0(\psi_t^{-1})^*)f.\end{aligned}$$

Let us suppose that $1\le n\le m$; the case $m \le n$ is similar. If $N < n$, all summands are 0. If $N = n$, we have by claim 4

$$A_Nf = (\partial_t^n|_0\varphi_t^*)f + (\partial_t^n|_0\psi_t^*)f + (\partial_t^n|_0(\varphi_t^{-1})^*)f + (\partial_t^n|_0(\psi_t^{-1})^*)f = 0.$$

If $n < N \le m$, we have, using again claim 4:

$$\begin{aligned}A_Nf &= \sum_{j+\ell=N}\frac{N!}{j!\ell!}(\partial_t^j|_0\psi_t^*)(\partial_t^\ell|_0(\psi_t^{-1})^*)f + \delta_N^m\big((\partial_t^m|_0\varphi_t^*)f + (\partial_t^m|_0(\varphi_t^{-1})^*)f\big)\\ &= (\partial_t^N|_0(\psi_t^{-1}\circ\psi_t)^*)f + 0 = 0.\end{aligned}$$

Now we come to the difficult case $m,n < N \le m+n$.

$$\begin{aligned}A_Nf &= \partial_t^N|_0(\psi_t^{-1}\circ\varphi_t^{-1}\circ\psi_t)^*f + \tbinom{N}{m}(\partial_t^m|_0\varphi_t^*)(\partial_t^{N-m}|_0(\psi_t^{-1}\circ\varphi_t^{-1}\circ\psi_t)^*)f\\ &\quad + (\partial_t^N|_0\varphi_t^*)f, \qquad (6)\end{aligned}$$

by claim 3, since all other terms vanish; see (8) below. By claim 3 again we get:

$$\begin{aligned}&\partial_t^N|_0(\psi_t^{-1}\circ\varphi_t^{-1}\circ\psi_t)^*f\\ &\qquad = \sum_{j+k+\ell=N}\frac{N!}{j!k!\ell!}(\partial_t^j|_0\psi_t^*)(\partial_t^k|_0(\varphi_t^{-1})^*)(\partial_t^\ell|_0(\psi_t^{-1})^*)f\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j+\ell=N} \tbinom{N}{j}(\partial_t^j|_0 \psi_t^*)(\partial_t^\ell|_0 (\psi_t^{-1})^*)f \\
&\quad + \tbinom{N}{m}(\partial_t^{N-m}|_0 \psi_t^*)(\partial_t^m|_0 (\varphi_t^{-1})^*)f \\
&\quad + \tbinom{N}{m}(\partial_t^m|_0 (\varphi_t^{-1})^*)(\partial_t^{N-m}|_0 (\psi_t^{-1})^*)f + \partial_t^N|_0 (\varphi_t^{-1})^* f \\
&= 0 + \tbinom{N}{m}(\partial_t^{N-m}|_0 \psi_t^*) m! \mathcal{L}_{-X} f + \tbinom{N}{m} m! \mathcal{L}_{-X}(\partial_t^{N-m}|_0 (\psi_t^{-1})^*)f \\
&\quad + \partial_t^N|_0 (\varphi_t^{-1})^* f \\
&= \delta_{m+n}^N (m+n)!(\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X) f + \partial_t^N|_0 (\varphi_t^{-1})^* f \\
&= \delta_{m+n}^N (m+n)! \mathcal{L}_{[X,Y]} f + \partial_t^N|_0 (\varphi_t^{-1})^* f. \qquad (7)
\end{aligned}
$$

From the second expression in (7) one can also read off that

$$\partial_t^{N-m}|_0 (\psi_t^{-1} \circ \varphi_t^{-1} \circ \psi_t)^* f = \partial_t^{N-m}|_0 (\varphi_t^{-1})^* f. \qquad (8)$$

If we put (7) and (8) into (6), we get, using claims 3 and 4 again, the final result which proves claim 5 and the theorem:

$$
\begin{aligned}
A_N f &= \delta_{m+n}^N (m+n)! \mathcal{L}_{[X,Y]} f + \partial_t^N|_0 (\varphi_t^{-1})^* f \\
&\quad + \tbinom{N}{m}(\partial_t^m|_0 \varphi_t^*)(\partial_t^{N-m}|_0 (\varphi_t^{-1})^*)f + (\partial_t^N|_0 \varphi_t^*)f \\
&= \delta_{m+n}^N (m+n)! \mathcal{L}_{[X,Y]} f + \partial_t^N|_0 (\varphi_t^{-1} \circ \varphi_t)^* f \\
&= \delta_{m+n}^N (m+n)! \mathcal{L}_{[X,Y]} f + 0. \quad \square
\end{aligned}
$$

3.17. Theorem. *Let $X_1, \dots, X_m$ be vector fields on M defined in a neighborhood of a point $x \in M$ such that $X_1(x), \dots, X_m(x)$ are a basis for T_xM and $[X_i, X_j] = 0$ for all i, j.*

Then there is a chart (U, u) of M centered at x such that $X_i|U = \frac{\partial}{\partial u^i}$.

Proof. For small $t = (t^1, \dots, t^m) \in \mathbb{R}^m$ we put

$$f(t^1, \dots, t^m) = (\mathrm{Fl}_{t^1}^{X_1} \circ \dots \circ \mathrm{Fl}_{t^m}^{X_m})(x).$$

By (3.15) we may interchange the order of the flows arbitrarily. Therefore

$$\tfrac{\partial}{\partial t^i} f(t^1, \dots, t^m) = \tfrac{\partial}{\partial t^i}(\mathrm{Fl}_{t^i}^{X_i} \circ \mathrm{Fl}_{t^1}^{X_1} \circ \cdots)(x) = X_i((\mathrm{Fl}_{t^1}^{x_1} \circ \cdots)(x)).$$

So $T_0 f$ is invertible, f is a local diffeomorphism, and its inverse gives a chart with the desired properties. $\square$

3.18. The theorem of Frobenius. The next three subsections will be devoted to the theorem of Frobenius for distributions of constant rank. We will give a powerful generalization for distributions of nonconstant rank below in (3.21) – (3.28).

Let M be a manifold. By a *vector subbundle* E of TM of fiber dimension k we mean a subset $E \subset TM$ such that each $E_x := E \cap T_xM$ is a linear subspace of dimension k and such that for each $x \operatorname{im} M$ there are k vector fields defined on an open neighborhood of M with values in E and spanning E, called a *local frame* for E. Such an E is also called a smooth *distribution* of constant rank k. See section (8) for a thorough discussion of the notion of vector bundles. The space of all vector fields with values in E will be called $\Gamma(E)$.

The vector subbundle E of TM is called *integrable* or *involutive*, if for all $X, Y \in \Gamma(E)$ we have $[X, Y] \in \Gamma(E)$.

Local version of Frobenius's theorem. *Let $E \subset TM$ be an integrable vector subbundle of fiber dimension k of TM.*

Then for each $x \in M$ there exists a chart (U, u) of M centered at x with $u(U) = V \times W \subset \mathbb{R}^k \times \mathbb{R}^{m-k}$, such that $T(u^{-1}(V \times \{y\})) = E|(u^{-1}(V \times \{y\}))$ for each $y \in W$.

Proof. Let $x \in M$. We choose a chart (U, u) of M centered at x such that there exist k vector fields $X_1, \dots, X_k \in \Gamma(E)$ which form a frame of $E|U$. Then we have $X_i = \sum_{j=1}^m f_i^j \frac{\partial}{\partial u^j}$ for $f_i^j \in C^\infty(U)$. Then $f = (f_i^j)$ is a $(k \times m)$-matrix valued smooth function on U which has rank k on U. So some $(k \times k)$-submatrix, say the top one, is invertible at x and thus we may take U so small that this top $(k \times k)$-submatrix is invertible everywhere on U. Let $g = (g_i^j)$ be the inverse of this submatrix, so that the $(k \times m)$-matrix $f.g$ is given by

$$f.g = \begin{pmatrix} \mathbb{I}_k \\ * \end{pmatrix}.$$

We put

$$Y_i := \sum_{j=1}^k g_i^j X_j = \sum_{j=1}^k \sum_{l=1}^m g_i^j f_j^l \frac{\partial}{\partial u^l} = \frac{\partial}{\partial u^i} + \sum_{p \geq k+1} h_i^p \frac{\partial}{\partial u^p}. \tag{1}$$

We claim that $[Y_i, Y_j] = 0$ for all $1 \leq i, j \leq k$. Since E is integrable, we have $[Y_i, Y_j] = \sum_{l=1}^k c_{ij}^l Y_l$. But from (1) we conclude (using the coordinate formula in (3.4)) that $[Y_i, Y_j] = \sum_{p \geq k+1} a^p \frac{\partial}{\partial u^p}$. Again by (1) this implies that $c_{ij}^l = 0$ for all l, and the claim follows.

Now we consider an $(m-k)$-dimensional linear subspace W_1 in $\mathbb{R}^m$ which is transversal to the k vectors $T_x u.Y_i(x) \in T_0\mathbb{R}^m$ spanning $\mathbb{R}^k$, and we define $f : V \times W \to U$ by

$$f(t^1, \dots, t^k, y) := \left(\mathrm{Fl}_{t^1}^{Y_1} \circ \mathrm{Fl}_{t^2}^{Y_2} \circ \dots \circ \mathrm{Fl}_{t^k}^{Y_k}\right)(u^{-1}(y)),$$

where $t = (t^1, \dots, t^k) \in V$, a small neighborhood of 0 in $\mathbb{R}^k$, and where $y \in W$, a small neighborhood of 0 in W_1. By (3.15) we may interchange the order of the flows in the definition of f arbitrarily. Thus

$$\frac{\partial}{\partial t^i} f(t,y) = \frac{\partial}{\partial t^i}\left(\mathrm{Fl}^{Y_i}_{t^i} \circ \mathrm{Fl}^{Y_1}_{t^1} \circ \dots\right)(u^{-1}(y)) = Y_i(f(t,y)),$$

$$\frac{\partial}{\partial y^k} f(0,y) = \frac{\partial}{\partial y^k}(u^{-1})(y),$$

and so $T_0 f$ is invertible and the inverse of f on a suitable neighborhood of x gives us the required chart. □

3.19. Remark. Any charts $(U, u : U \to V \times W \subset \mathbb{R}^k \times \mathbb{R}^{m-k})$ as constructed in theorem (3.18) with V and W open balls is called a *distinguished chart* for E. The submanifolds $u^{-1}(V \times \{y\})$ are called *plaques*. Two plaques of different distinguished charts intersect in open subsets in both plaques or not at all: This follows immediately by flowing a point in the intersection into both plaques with the same construction as in the proof of (3.18). Thus an atlas of distinguished charts on M has chart change mappings which respect the submersion $\mathbb{R}^k \times \mathbb{R}^{m-k} \to \mathbb{R}^{m-k}$ (the plaque structure on M). Such an atlas (or the equivalence class of such atlases) is called the *foliation corresponding to the integrable vector subbundle $E \subset TM$.*

3.20. Global version of Frobenius's theorem. *Let $E \subsetneq TM$ be an integrable vector subbundle of TM. Then, using the restrictions of distinguished charts to plaques as charts, we get a new structure of a smooth manifold on M, which we denote by M_E. If $E \neq TM$, the topology of M_E is finer than that of M, M_E has uncountably many connected components called the leaves of the foliation, and the identity induces a bijective immersion $M_E \to M$. Each leaf L is a second countable initial submanifold of M, and it is a maximal integrable submanifold of M for E in the sense that $T_x L = E_x$ for each $x \in L$.*

Proof. Let $(U_\alpha, u_\alpha : U_\alpha \to V_\alpha \times W_\alpha \subseteq \mathbb{R}^k \times \mathbb{R}^{m-k})$ be an atlas of distinguished charts corresponding to the integrable vector subbundle $E \subset TM$, as given by theorem (3.18). Let us now use for each plaque the homeomorphisms $\mathrm{pr}_1 \circ u_\alpha|(u_\alpha^{-1}(V_\alpha \times \{y\})) : u_\alpha^{-1}(V_\alpha \times \{y\}) \to V_\alpha \subset \mathbb{R}^{m-k}$ as charts; then we describe on M a new smooth manifold structure M_E with finer topology which however has uncountably many connected components, and the identity on M induces a bijective immersion $M_E \to M$. The connected components of M_E are called the *leaves of the foliation*.

In order to check the rest of the assertions made in the theorem, let us construct the unique leaf L through an arbitrary point $x \in M$: choose a plaque containing x and take the union with any plaque meeting the first one, and keep going. Now choose $y \in L$ and a curve $c : [0,1] \to L$ with $c(0) = x$ and $c(1) = y$. Then there are finitely many distinguished charts

$(U_1, u_1), \dots, (U_n, u_n)$ and $a_1, \dots, a_n \in \mathbb{R}^{m-k}$ such that $x \in u_1^{-1}(V_1 \times \{a_1\})$, $y \in u_n^{-1}(V_n \times \{a_n\})$ and such that for each i

$$(1) \qquad u_i^{-1}(V_i \times \{a_i\}) \cap u_{i+1}^{-1}(V_{i+1} \times \{a_{i+1}\}) \neq \emptyset.$$

Given u_i, u_{i+1}, and a_i, there are only countably many points a_{i+1} such that (1) holds: If not, then we get a cover of the the separable submanifold $u_i^{-1}(V_i \times \{a_i\}) \cap U_{i+1}$ by uncountably many pairwise disjoint open sets of the form given in (1), which contradicts separability.

Finally, since (each component of) M is a Lindelöf space, any distinguished atlas contains a countable subatlas. So each leaf is the union of at most countably many plaques. The rest is clear. □

3.21. Singular distributions. Let M be a manifold. Suppose that for each $x \in M$ we are given a vector subspace E_x of T_xM. The disjoint union $E = \bigsqcup_{x \in M} E_x$ is called a *(singular) distribution* on M. We do not suppose that the dimension of E_x is locally constant in x.

Let $\mathfrak{X}_{loc}(M)$ denote the set of all locally defined smooth vector fields on M, i.e., $\mathfrak{X}_{loc}(M) = \bigcup \mathfrak{X}(U)$, where U runs through all open sets in M. Furthermore let $\mathfrak{X}_E$ denote the set of all local vector fields $X \in \mathfrak{X}_{loc}(M)$ with $X(x) \in E_x$ whenever defined. We say that a subset $\mathcal{V} \subset \mathfrak{X}_E$ *spans* E if for each $x \in M$ the vector space E_x is the linear hull of the set $\{X(x) : X \in \mathcal{V}\}$. We say that E is a *smooth distribution* if $\mathfrak{X}_E$ spans E. Note that every subset $\mathcal{W} \subset \mathfrak{X}_{loc}(M)$ spans a distribution denoted by $E(\mathcal{W})$, which is obviously smooth (the linear span of the empty set is the vector space 0). From now on we will consider only smooth distributions.

An *integral manifold* of a smooth distribution E is a connected immersed submanifold (N, i) (see (2.9)) such that $T_x i(T_x N) = E_{i(x)}$ for all $x \in N$. We will see in theorem (3.25) below that any integral manifold is in fact an initial submanifold of M (see (2.13)), so that we need not specify the injective immersion i. An integral manifold of E is called *maximal* if it is not contained in any strictly larger integral manifold of E.

3.22. Lemma. *Let E be a smooth distribution on M. Then we have:*

(1) *If (N, i) is an integral manifold of E and $X \in \mathfrak{X}_E$, then i^*X makes sense and is an element of $\mathfrak{X}_{loc}(N)$, which is $i|i^{-1}(U_X)$-related to X, where $U_X \subset M$ is the open domain of X.*

(2) *If (N_j, i_j) are integral manifolds of E for $j = 1, 2$, then $i_1^{-1}(i_1(N_1) \cap i_2(N_2))$ and $i_2^{-1}(i_1(N_1) \cap i_2(N_2))$ are open subsets in N_1 and N_2, respectively; furthermore $i_2^{-1} \circ i_1$ is a diffeomorphism between them.*

(3) *If $x \in M$ is contained in some integral submanifold of E, then it is contained in a unique maximal one.*

Proof. (1) Let U_X be the open domain of $X \in \mathfrak{X}_E$. If $i(x) \in U_X$ for $x \in N$, we have $X(i(x)) \in E_{i(x)} = T_x i(T_x N)$, so $i^*X(x) := ((T_x i)^{-1} \circ X \circ i)(x)$ makes sense. The vector field i^*X is clearly defined on an open subset of N and is smooth.

(2) Let $X \in \mathfrak{X}_E$. Then $i_j^*X \in \mathfrak{X}_{loc}(N_j)$ and is i_j-related to X. So by lemma (3.14) for $j = 1, 2$ we have

$$i_j \circ \mathrm{Fl}_t^{i_j^* X} = Fl_t^X \circ i_j.$$

Now choose $x_j \in N_j$ such that $i_1(x_1) = i_2(x_2) = x_0 \in M$ and choose vector fields $X_1, \dots, X_n \in \mathfrak{X}_E$ such that $(X_1(x_0), \dots, X_n(x_0))$ is a basis of E_{x_0}. Then

$$f_j(t^1, \dots, t^n) := (\mathrm{Fl}_{t^1}^{i_j^* X_1} \circ \dots \circ \mathrm{Fl}_{t^n}^{i_j^* X_n})(x_j)$$

is a smooth local mapping $\mathbb{R}^n \to N_j$ defined near zero. Since obviously $\frac{\partial}{\partial t^k}|_0 f_j = i_j^* X_k(x_j)$ for $j = 1, 2$, we see that f_j is a diffeomorphism near 0. Finally we have

$$\begin{aligned}(i_2^{-1} \circ i_1 \circ f_1)(t^1, \dots, t^n) &= (i_2^{-1} \circ i_1 \circ \mathrm{Fl}_{t^1}^{i_1^* X_1} \circ \dots \circ \mathrm{Fl}_{t^n}^{i_1^* X_n})(x_1) \\ &= (i_2^{-1} \circ \mathrm{Fl}_{t^1}^{X_1} \circ \dots \circ \mathrm{Fl}_{t^n}^{X_n} \circ i_1)(x_1) \\ &= (\mathrm{Fl}_{t^1}^{i_2^* X_1} \circ \dots \circ \mathrm{Fl}_{t^n}^{i_2^* X_n} \circ i_2^{-1} \circ i_1)(x_1) \\ &= f_2(t^1, \dots, t^n).\end{aligned}$$

So $i_2^{-1} \circ i_1$ is a diffeomorphism, as required.

(3) Let N be the union of all integral manifolds containing x. Choose the union of all the atlases of these integral manifolds as atlas for N, which is a smooth atlas for N by (2). Note that a connected immersed submanifold of a separable manifold is automatically separable (since it carries a Riemann metric). □

3.23. Integrable singular distributions and singular foliations. A smooth singular distribution E on a manifold M is called *integrable* if each point of M is contained in some integral manifold of E. By (3.22.3) each point is then contained in a unique maximal integral manifold, so the maximal integral manifolds form a partition of M. This partition is called the *(singular) foliation* of M induced by the integrable (singular) distribution E, and each maximal integral manifold is called a *leaf* of this foliation. If $X \in \mathfrak{X}_E$, then by (3.22.1) the integral curve $t \mapsto \mathrm{Fl}^X(t, x)$ of X through $x \in M$ stays in the leaf through x.

Let us now consider an arbitrary subset $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$. We say that $\mathcal{V}$ is *stable* if for all $X, Y \in \mathcal{V}$ and for all t for which it is defined the local vector field $(\mathrm{Fl}_t^X)^* Y$ is again an element of $\mathcal{V}$.

If $\mathcal{W} \subset \mathfrak{X}_{loc}(M)$ is an arbitrary subset, we call $\mathcal{S}(\mathcal{W})$ the set of all local vector fields of the form $(\mathrm{Fl}^{X_1}_{t_1} \circ \dots \circ \mathrm{Fl}^{X_k}_{t_k})^* Y$ for $X_i, Y \in \mathcal{W}$. By lemma (3.14) the flow of this vector field is

$$\mathrm{Fl}((\mathrm{Fl}^{X_1}_{t_1} \circ \dots \circ \mathrm{Fl}^{X_k}_{t_k})^* Y, t) = \mathrm{Fl}^{X_k}_{-t_k} \circ \dots \circ \mathrm{Fl}^{X_1}_{-t_1} \circ \mathrm{Fl}^{Y}_{t} \circ \mathrm{Fl}^{X_1}_{t_1} \circ \dots \circ \mathrm{Fl}^{X_k}_{t_k},$$

so $\mathcal{S}(\mathcal{W})$ is the minimal stable set of local vector fields which contains $\mathcal{W}$.

Now let F be an arbitrary distribution. A local vector field $X \in \mathfrak{X}_{\mathrm{loc}}(M)$ is called an *infinitesimal automorphism* of F if $T_x(\mathrm{Fl}^X_t)(F_x) \subset F_{\mathrm{Fl}^X(t,x)}$ whenever defined. We denote by $\mathrm{aut}(F)$ the set of all infinitesimal automorphisms of F. By arguments given just above, $\mathrm{aut}(F)$ is stable.

3.24. Lemma. *Let E be a smooth distribution on a manifold M. Then the following conditions are equivalent:*

(1) *E is integrable.*
(2) *$\mathfrak{X}_E$ is stable.*
(3) *There exists a subset $\mathcal{W} \subset \mathfrak{X}_{loc}(M)$ such that $\mathcal{S}(\mathcal{W})$ spans E.*
(4) *$aut(E) \cap \mathfrak{X}_E$ spans E.*

Proof. (1) $\Longrightarrow$ (2) Let $X \in \mathfrak{X}_E$ and let L be the leaf through $x \in M$, with $i : L \to M$ the inclusion. Then $\mathrm{Fl}^X_{-t} \circ i = i \circ \mathrm{Fl}^{i^*X}_{-t}$ by lemma (3.14), so we have

$$\begin{aligned} T_x(\mathrm{Fl}^X_{-t})(E_x) &= T(\mathrm{Fl}^X_{-t}).T_x i.T_x L = T(\mathrm{Fl}^X_{-t} \circ i).T_x L \\ &= Ti.T_x(\mathrm{Fl}^{i^*X}_{-t}).T_x L \\ &= Ti.T_{Fl^{i^*X}(-t,x)} L = E_{Fl^X(-t,x)}. \end{aligned}$$

This implies that $(\mathrm{Fl}^X_t)^* Y \in \mathfrak{X}_E$ for any $Y \in \mathfrak{X}_E$.

(2) $\Longrightarrow$ (4) In fact (2) says that $\mathfrak{X}_E \subset aut(E)$.

(4) $\Longrightarrow$ (3) We can choose $\mathcal{W} = aut(E) \cap \mathfrak{X}_E$: For $X, Y \in \mathcal{W}$ we have $(\mathrm{Fl}^X_t)^* Y \in \mathfrak{X}_E$; so $\mathcal{W} \subset \mathcal{S}(\mathcal{W}) \subset \mathfrak{X}_E$ and E is spanned by $\mathcal{W}$.

(3) $\Longrightarrow$ (1) We have to show that each point $x \in M$ is contained in some integral submanifold for the distribution E. Since $\mathcal{S}(\mathcal{W})$ spans E and is stable, we have

$$T(\mathrm{Fl}^X_t).E_x = E_{\mathrm{Fl}^X(t,x)} \tag{5}$$

for each $X \in \mathcal{S}(\mathcal{W})$. Let $\dim E_x = n$. There are $X_1, \dots, X_n \in \mathcal{S}(\mathcal{W})$ such that $X_1(x), \dots, X_n(x)$ is a basis of E_x, since E is smooth. As in the proof of (3.22.2) we consider the mapping

$$f(t^1, \dots, t^n) := (\mathrm{Fl}^{X_1}_{t^1} \circ \dots \circ \mathrm{Fl}^{X_n}_{t^n})(x),$$

defined and smooth near 0 in $\mathbb{R}^n$. Since the rank of f at 0 is n, the image under f of a small open neighborhood of 0 is a submanifold N of M. We claim that N is an integral manifold of E. The tangent space $T_{f(t^1,\dots,t^n)}N$ is linearly generated by

$$\begin{aligned}\tfrac{\partial}{\partial t^k}(\mathrm{Fl}^{X_1}_{t^1}\circ\dots\circ\mathrm{Fl}^{X_n}_{t^n})(x) &= T(\mathrm{Fl}^{X_1}_{t^1}\circ\dots\circ\mathrm{Fl}^{X_{k-1}}_{t^{k-1}})X_k((\mathrm{Fl}^{X_k}_{t^k}\circ\dots\circ\mathrm{Fl}^{X_n}_{t^n})(x))\\ &= ((\mathrm{Fl}^{X_1}_{-t^1})^*\cdots(\mathrm{Fl}^{X_{k-1}}_{-t^{k-1}})^*X_k)(f(t^1,\dots,t^n)).\end{aligned}$$

Since $\mathcal{S}(\mathcal{W})$ is stable, these vectors lie in $E_{f(t)}$. From the form of f and from (5) we see that $\dim E_{f(t)} = \dim E_x$, so these vectors even span $E_{f(t)}$ and we have $T_{f(t)}N = E_{f(t)}$ as required. □

3.25. Theorem (Local structure of singular foliations). *Let E be an integrable (singular) distribution of a manifold M. Then for each $x \in M$ there exist a chart (U,u) with $u(U) = \{y \in \mathbb{R}^m : |y^i| < \varepsilon$ for all $i\}$ for some $\varepsilon > 0$ and a countable subset $A \subset \mathbb{R}^{m-n}$, such that for the leaf L through x we have*

$$u(U \cap L) = \{y \in u(U) : (y^{n+1},\dots,y^m) \in A\}.$$

Each leaf is an initial submanifold.

If furthermore the distribution E has locally constant rank, this property holds for each leaf meeting U with the same n.

This chart (U,u) is called a *distinguished chart* for the (singular) distribution or the (singular) foliation. A connected component of $U\cap L$ is called a *plaque*.

Proof. Let L be the leaf through x, $\dim L = n$. Let $X_1,\dots,X_n \in \mathfrak{X}_E$ be local vector fields such that $X_1(x),\dots,X_n(x)$ is a basis of E_x. We choose a chart (V,v) centered at x on M such that the vectors

$$X_1(x),\dots,X_n(x), \tfrac{\partial}{\partial v^{n+1}}|_x,\dots,\tfrac{\partial}{\partial v^m}|_x$$

form a basis of T_xM. Then

$$f(t^1,\dots,t^m) = (\mathrm{Fl}^{X_1}_{t^1}\circ\dots\circ\mathrm{Fl}^{X_n}_{t^n})(v^{-1}(0,\dots,0,t^{n+1},\dots,t^m))$$

is a diffeomorphism from a neighborhood of 0 in $\mathbb{R}^m$ onto a neighborhood of x in M. Let (U,u) be the chart given by f^{-1}, suitably restricted. We have

$$y \in L \iff (\mathrm{Fl}^{X_1}_{t^1}\circ\dots\circ\mathrm{Fl}^{X_n}_{t^n})(y) \in L$$

for all y and all $t^1,\dots,t^n$ for which both expressions make sense. So we have

$$f(t^1,\dots,t^m) \in L \iff f(0,\dots,0,t^{n+1},\dots,t^m) \in L,$$

and consequently $L\cap U$ is the disjoint union of connected sets of the form $\{y \in U : (u^{n+1}(y),\dots,u^m(y)) = \text{constant}\}$. Since L is a connected immersive submanifold of M, it is second countable and only a countable set of constants can appear in the description of $u(L\cap U)$ given above. From this description it is clear that L is an initial submanifold (2.13) since $u(C_x(L\cap U)) = u(U)\cap(\mathbb{R}^n \times 0)$.

The argument given above is valid for any leaf of dimension n meeting U, so also the assertion for an integrable distribution of constant rank follows. $\square$

3.26. Involutive singular distributions. A subset $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$ is called *involutive* if $[X, Y] \in \mathcal{V}$ for all $X, Y \in \mathcal{V}$. Here $[X, Y]$ is defined on the intersection of the domains of X and Y.

A smooth distribution E on M is called *involutive* if there exists an involutive subset $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$ spanning E.

For an arbitrary subset $\mathcal{W} \subset \mathfrak{X}_{loc}(M)$ let $\mathcal{L}(\mathcal{W})$ be the set consisting of all local vector fields on M which can be written as finite expressions using Lie brackets and starting from elements of $\mathcal{W}$. Clearly $\mathcal{L}(\mathcal{W})$ is the smallest involutive subset of $\mathfrak{X}_{loc}(M)$ which contains $\mathcal{W}$.

3.27. Lemma. *For each subset $\mathcal{W} \subset \mathfrak{X}_{loc}(M)$ we have*

$$E(\mathcal{W}) \subset E(\mathcal{L}(\mathcal{W})) \subset E(\mathcal{S}(\mathcal{W})).$$

In particular we have $E(\mathcal{S}(\mathcal{W})) = E(\mathcal{L}(\mathcal{S}(\mathcal{W})))$.

Proof. We will show that for $X, Y \in \mathcal{W}$ we have $[X, Y] \in \mathfrak{X}_{E(\mathcal{S}(\mathcal{W}))}$, for then by induction we get $\mathcal{L}(\mathcal{W}) \subset \mathfrak{X}_{E(\mathcal{S}(\mathcal{W}))}$ and $E(\mathcal{L}(\mathcal{W})) \subset E(\mathcal{S}(\mathcal{W}))$.

Let $x \in M$; since by (3.24) $E(\mathcal{S}(\mathcal{W}))$ is integrable, we can choose the leaf L through x, with the inclusion i. Then i^*X is i-related to X and i^*Y is i-related to Y; thus by (3.10) the local vector field $[i^*X, i^*Y] \in \mathfrak{X}_{loc}(L)$ is i-related to $[X, Y]$, and $[X, Y](x) \in E(\mathcal{S}(\mathcal{W}))_x$, as required. $\square$

3.28. Theorem. *Let $\mathcal{V} \subset \mathfrak{X}_{loc}(M)$ be an involutive subset. Then the distribution $E(\mathcal{V})$ spanned by $\mathcal{V}$ is integrable under each of the following conditions.*

(1) *M is real analytic and $\mathcal{V}$ consists of real analytic vector fields.*

(2) *The dimension of $E(\mathcal{V})$ is constant along flow lines of vector fields in $\mathcal{V}$.*

Proof. (1) For $X, Y \in \mathcal{V}$ we have $\frac{d}{dt}(\mathrm{Fl}^X_t)^*Y = (\mathrm{Fl}^X_t)^*\mathcal{L}_X Y$; consequently $\frac{d^k}{dt^k}(\mathrm{Fl}^X_t)^*Y = (\mathrm{Fl}^X_t)^*(\mathcal{L}_X)^k Y$, and since everything is real analytic, we get for $x \in M$ and small t

$$(\mathrm{Fl}^X_t)^*Y(x) = \sum_{k\ge 0} \frac{t^k}{k!}\frac{d^k}{dt^k}|_0(\mathrm{Fl}^X_t)^*Y(x) = \sum_{k\ge 0}\frac{t^k}{k!}(\mathcal{L}_X)^k Y(x).$$

Since $\mathcal{V}$ is involutive, all $(\mathcal{L}_X)^k Y \in \mathcal{V}$. Therefore we get $(\mathrm{Fl}^X_t)^*Y(x) \in E(\mathcal{V})_x$ for small t. By the flow property of Fl^X the set of all t satisfying $(\mathrm{Fl}^X_t)^*Y(x) \in E(\mathcal{V})_x$ is open and closed, so it follows that (3.24.2) is satisfied and thus $E(\mathcal{V})$ is integrable.

(2) We choose $X_1, \dots, X_n \in \mathcal{V}$ such that $X_1(x), \dots, X_n(x)$ is a basis of $E(\mathcal{V})_x$. For any $X \in \mathcal{V}$, by hypothesis, $E(\mathcal{V})_{\mathrm{Fl}^X(t,x)}$ has also dimension n

and admits the vectors $X_1(\mathrm{Fl}^X(t,x)),\dots,X_n(\mathrm{Fl}^X(t,x))$ as basis, for small t. So there are smooth functions $f_{ij}(t)$ such that

$$[X,X_i](\mathrm{Fl}^X(t,x)) = \sum_{j=1}^n f_{ij}(t)X_j(\mathrm{Fl}^X(t,x)).$$

Therefore,

$$\begin{aligned}\tfrac{d}{dt}T(\mathrm{Fl}^X_{-t}).X_i(\mathrm{Fl}^X(t,x)) &= T(\mathrm{Fl}^X_{-t}).[X,X_i](\mathrm{Fl}^X(t,x))\\ &= \sum_{j=1}^n f_{ij}(t)T(\mathrm{Fl}^X_{-t}).X_j(\mathrm{Fl}^X(t,x)).\end{aligned}$$

So the T_xM-valued functions $g_i(t) = T(\mathrm{Fl}^X_{-t}).X_i(\mathrm{Fl}^X(t,x))$ satisfy the linear ordinary differential equation $\frac{d}{dt}g_i(t) = \sum_{j=1}^n f_{ij}(t)g_j(t)$ and have initial values in the linear subspace $E(\mathcal{V})_x$, so they have values in it for all small t. Therefore $T(\mathrm{Fl}^X_{-t})E(\mathcal{V})_{\mathrm{Fl}^X(t,x)} \subset E(\mathcal{V})_x$ for small t. Using compact time intervals and the flow property, one sees that condition (3.24.2) is satisfied and $E(\mathcal{V})$ is integrable. □

3.29. Examples. (1) The singular distribution spanned by $\mathcal{W} \subset \mathfrak{X}_{loc}(\mathbb{R}^2)$ is involutive, but not integrable, where $\mathcal{W}$ consists of all global vector fields with support in $\mathbb{R}^2 \setminus \{0\}$ and the field $\frac{\partial}{\partial x^1}$; the leaf through 0 should have dimension 1 at 0 and dimension 2 elsewhere.

(2) Let $f : \mathbb{R} \to \mathbb{R}$ be a smooth function with $f(x^1) = 0$ for $x^1 \le 0$ and $f(x^1) > 0$ for $x^1 > 0$. Then the singular distribution on $\mathbb{R}^2$ spanned by the two vector fields $X(x^1,x^2) = \frac{\partial}{\partial x^1}$ and $Y(x^1,x^2) = f(x^1)\frac{\partial}{\partial x^2}$ is involutive, but not integrable. Any leaf should pass $(0,x^2)$ tangentially to $\frac{\partial}{\partial x^1}$, should have dimension 1 for $x^1 \le 0$ and should have dimension 2 for $x^1 > 0$.

3.30. By a *time dependent vector field* on a manifold M we mean a smooth mapping $X : J \times M \to TM$ with $\pi_M \circ X = \mathrm{pr}_2$, where J is an open interval. An integral curve of X is a smooth curve $c : I \to M$ with $\dot c(t) = X(t,c(t))$ for all $t \in I$, where I is a subinterval of J.

There is an associated vector field $\bar X \in \mathfrak{X}(J \times M)$, given by $\bar X(t,x) = (\frac{\partial}{\partial t}, X(t,x)) \in T_t\mathbb{R} \times T_xM$.

By the *evolution operator* of X we mean the mapping $\Phi^X : J \times J \times M \to M$, defined in a maximal open neighborhood of $\Delta_J \times M$ (where Δ_J is the diagonal of J) and satisfying the differential equation

$$\begin{cases} \frac{d}{dt}\Phi^X(t,s,x) = X(t,\Phi^X(t,s,x)) \\ \Phi^X(s,s,x) = x. \end{cases}$$

It is easily seen that $(t, \Phi^X(t,s,x)) = \mathrm{Fl}^{\bar X}(t-s,(s,x))$, so the maximally defined evolution operator exists and is unique, and it satisfies

$$\Phi^X_{t,s} = \Phi^X_{t,r} \circ \Phi^X_{r,s}, \qquad \text{where } \Phi^X_{t,s}(x) = \Phi(t,s,x),$$

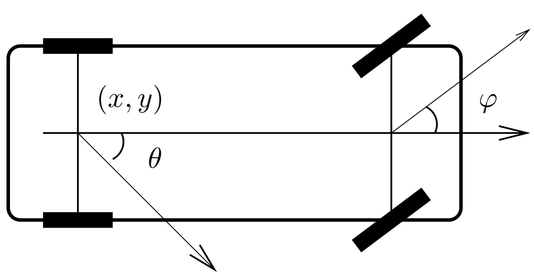
whenever one side makes sense (with the restrictions of (3.7)).

Examples and Exercises

3.31. Compute the flow of the vector field $\xi_1(x,y) := y\frac{\partial}{\partial x}$ in $\mathbb{R}^2$. Is it a global flow? Answer the same questions for $\xi_2(x,y) := \frac{x^2}{2}\frac{\partial}{\partial y}$. Now compute $[\xi_1,\xi_2]$ and investigate its flow. This time it is not global! In fact, $Fl_t^{[\xi_1,\xi_2]}(x,y) = \left(\frac{2x}{2+xt}, \frac{y}{4}(tx+2)^2\right)$. Investigate the flow of $\xi_1+\xi_2$. It is not global either! Thus the set of complete vector fields on $\mathbb{R}^2$ is neither a vector space nor closed under the Lie bracket.

3.32. Driving a car. The phase space consists of all $(x,y,\vartheta,\varphi) \in \mathbb{R}^2 \times S^1 \times (-\pi/4, \pi/4)$, where

(x,y)	is the position of the midpoint of the rear axle,
ϑ	is the direction of the car axle,
ϕ	is the steering angle of the front wheels.



There are two 'control' vector fields:

$$\begin{aligned} \text{steer } &= \frac{\partial}{\partial \phi}, \\ \text{drive } &= \cos(\vartheta)\frac{\partial}{\partial x} + \sin(\vartheta)\frac{\partial}{\partial y} + \tan(\phi)\frac{1}{l}\frac{\partial}{\partial \vartheta} \text{ (why?).} \end{aligned}$$

Compute $[\text{steer}, \text{drive}] =: \text{park}$ (why?) and $[\text{drive}, \text{park}]$, and interpret the results. Is it not convenient that the two control vector fields do not span an integrable distribution?

3.33. Describe the Lie algebra of all vector fields on S^1 in terms of Fourier expansion. This is nearly (up to a central extension) the Virasoro algebra of theoretical physics.