Fiorella Barone and Paolo Facchi

Hamiltonian Systems

63/2007

Hamiltonian Systems

Fiorella Barone

Dipartimento di Matematica, Università di Bari, Italy

Paolo Facchi

Dipartimento di Matematica, Università di Bari, Italy

Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Italy

This work describes the fundamental principles and methods of Hamiltonian systems. In the first part the necessary apparatus of differential geometry is introduced starting from manifolds up to Cartan's calculus. The second part is devoted to the study of Hamiltonian systems and, in particular, to reduction theory, Hamilton-Jacobi equation and perturbation theory.

Rapporto n. *63/2007*

I

Differential Geometry

II
Hamiltonian systems

I.

Differential Geometry

1 Smooth manifolds and smooth mappings

1.1 Smooth manifolds

Let M be a set and m a positive integer. Any couple $(\mathcal{U}, \xi)$, with

$$\xi : \mathcal{U} \subset M \longrightarrow \mathbb{R}^m$$

injective mapping, is called an *m-chart* on M (*global*, in the case $\mathcal{U} = M$). Compositions

$$x^i : \mathcal{U} \xrightarrow{\ \xi\ } \mathbb{R}^m \xrightarrow{\ pr^i\ } \mathbb{R}$$

for $(i = 1, \ldots, m)$, denote the *coordinate functions* of ξ. For any point $x \in \mathcal{U}$,

$$\xi(x) = \big(x^1(x), \ldots, x^m(x)\big)$$

is the m-tuple of *coordinates* of x in ξ. The bijection induced by ξ onto its own image $\xi(\mathcal{U})$ is still denoted by $\xi : \mathcal{U} \to \xi(\mathcal{U})$ and then the inverse bijection by $\xi^{-1} : \xi(\mathcal{U}) \to \mathcal{U}$.
Let $(\mathcal{U}, \xi)$ and $(\mathcal{V}, \eta)$ be two m-charts on M. They are said to be C^∞*–related* to each other if $\mathcal{U} \cap \mathcal{V} = \emptyset$ or, when $\mathcal{U} \cap \mathcal{V} \neq \emptyset$, if their *transition functions*

$$\eta \circ \xi^{-1} : \xi(\mathcal{U} \cap \mathcal{V}) \longrightarrow \eta(\mathcal{U} \cap \mathcal{V})$$
$$\xi \circ \eta^{-1} : \eta(\mathcal{V} \cap \mathcal{U}) \longrightarrow \xi(\mathcal{U} \cap \mathcal{V})$$

are C^∞ (which implies that both $\xi(\mathcal{U} \cap \mathcal{V})$ and $\eta(\mathcal{V} \cap \mathcal{U})$ are open subsets of $\mathbb{R}^m$). Notice that an m-chart is C^∞–related to itself if, and only if, its image is an open subset of $\mathbb{R}^m$. In the sequel we will denote a chart $(\mathcal{U}, \xi)$ simply by ξ if no ambiguity occurs.

A collection $\mathcal{A}$ of (m-)charts is said to be an (m-dimensional) *atlas* on M if
$(\mathcal{A}_1)$ the domains of the charts belonging to $\mathcal{A}$ are a covering of M.
An (m-dimensional) atlas $\mathcal{A}$ is said to be C^∞ differentiable if
$(\mathcal{A}_2)$ for each $\xi \in \mathcal{A}$, ξ is C^∞–related to every chart of $\mathcal{A}$.
An (m-dimensional) C^∞ atlas $\mathcal{A}$ is said to be complete if
$(\mathcal{A}_3)$ any (m-)chart C^∞–related to every chart of $\mathcal{A}$, belongs to $\mathcal{A}$.

1.1.1 **Proposition.** *Each (m-dimensional) C^∞ atlas $\mathcal{A}$ on M is contained in just one complete (m-dimensional) C^∞ atlas $\mathcal{C}$, given by*

$$\mathcal{C} = \{\, \xi \mid \xi \text{ is an } m\text{-chart on } M,\ C^\infty\text{–related to every chart of } \mathcal{A} \,\}.$$

Proof. Let us consider the above collection $\mathcal{C}$ of m-charts. From the second property of atlas, we deduce that $\mathcal{A} \subset \mathcal{C}$. This also implies that $\mathcal{C}$ satisfies covering property $(\mathcal{A}_1)$. Now, let $(\mathcal{U}_1, \xi_1)$ and

$(\mathcal{U}_2, \xi_2)$ be any two charts of $\mathcal{C}$ with $\mathcal{U}_1 \cap \mathcal{U}_2 \neq \emptyset$. In order to prove their C^∞–relatedness, choose a chart $(\eta, \mathcal{V}) \in \mathcal{A}$ such that $\mathcal{U}_1 \cap \mathcal{U}_2 \cap \mathcal{V} \neq \emptyset$ and consider the restrictions

$$(\xi_2 \circ \xi_1^{-1})|_{\xi_1(\mathcal{U}_1 \cap \mathcal{U}_2 \cap \mathcal{V})} = (\xi_2 \circ \eta^{-1}) \circ (\eta \circ \xi_1^{-1})$$

$$(\xi_1 \circ \xi_2^{-1})|_{\xi_2(\mathcal{U}_1 \cap \mathcal{U}_2 \cap \mathcal{V})} = (\xi_1 \circ \eta^{-1}) \circ (\eta \circ \xi_2^{-1})$$

Since the transition functions $\eta \circ \xi_1^{-1}$ and $\xi_2 \circ \eta^{-1}$ are C^∞, the above restrictions are C^∞ and then, owing to the arbitrarity of η, the original functions too. This shows that $\mathcal{C}$ satisfies C^∞ property $(\mathcal{A}_2)$. Moreover, if ξ is an m-chart C^∞-related to every chart of $\mathcal{C}$, ξ is, in particular, C^∞–related to every chart of $\mathcal{A} \subset \mathcal{C}$ and then $\xi \in \mathcal{C}$. This shows that $\mathcal{C}$ satisfies completeness property $(\mathcal{A}_3)$. As to the uniqueness of $\mathcal{C}$, let us consider any complete m-dimensional C^∞ atlas $\mathcal{C}'$ containing $\mathcal{A}$. A chart $\xi \in \mathcal{C}'$ is C^∞–related to every chart of $\mathcal{A} \subset \mathcal{C}'$ and then $\xi \in \mathcal{C}$. As a consequence, each chart $\eta \in \mathcal{C}$ is C^∞–related to every chart of $\mathcal{C}' \subset \mathcal{C}$ and then, for completeness reasons, $\eta \in \mathcal{C}'$. This shows that $\mathcal{C}' = \mathcal{C}$. ∎

A complete m-dimensional C^∞ atlas is also called an m-dimensional *differential structure* on M. Distinct atlases whose charts are C^∞–related to one other, all determine – owing to the above proposition – the same differential structure on M.

1.1.2 **Example.** Let V be an m-dimensional (real) vector space. Any linear basis of V defines a linear isomorphism of V onto $\mathbb{R}^m$, which is an m-dimensional C^∞ atlas; all of these atlases determine the same differential structure on V. In the particular case $V = \mathbb{R}^m$, such a differential structure is determined by the distinguished global chart $\mathrm{id}_{\mathbb{R}^m}$, which corresponds to the canonical basis $(\delta_i)_{i=1,\ldots,m}$ with $\delta_i := (\delta_i^h)_{h=1,\ldots,m}$ (*Kronecker symbols*). ∎

A set M equipped with an m-dimensional differential structure $\mathcal{C}$ is called an m-dimensional *smooth manifold*. All the charts of $\mathcal{C}$ are called *admissible charts* on M and their domains *coordinate domains* on M. Given a point $x \in M$, we shall say that $(\mathcal{U}, \xi)$ is a chart at x if $x \in \mathcal{U}$.
Coordinate domains set up a basis of a topology on M. To show that, we need the following

1.1.3 **Lemma.** *Let $(\mathcal{U}_1, \xi_1)$ and $(\mathcal{U}_2, \xi_2)$ be admissible charts; restrictions $(\mathcal{U}_1 \cap \mathcal{U}_2, \xi_i|_{\mathcal{U}_1 \cap \mathcal{U}_2})$, $i = 1, 2$, are admissible charts, too.*

Proof. Il is sufficient to assume $\mathcal{U}_1 \cap \mathcal{U}_2 \neq \emptyset$ and consider the case $i = 1$. Let us put $\mathcal{U} = \mathcal{U}_1 \cap \mathcal{U}_2$ and $\xi = \xi_1|_{\mathcal{U}_1 \cap \mathcal{U}_2}$. We have to prove that $(\mathcal{U}, \xi) \in \mathcal{C}$. To this end consider any other chart $(\mathcal{V}, \eta) \in \mathcal{C}$ such that $\mathcal{U} \cap \mathcal{V} \neq \emptyset$. Owing to the C^∞–relatedness of ξ_1, ξ_2 and η, the sets

$$\xi(\mathcal{U} \cap \mathcal{V}) = \xi_1(\mathcal{U}_1 \cap \mathcal{V}) \cap \xi_1(\mathcal{U}_1 \cap \mathcal{U}_2)$$
$$\eta(\mathcal{U} \cap \mathcal{V}) = \eta(\mathcal{U}_1 \cap \mathcal{V}) \cap \eta(\mathcal{U}_2 \cap \mathcal{V})$$

are both open in $\mathbb{R}^m$. and then the transition functions

$$\eta \circ \xi^{-1} = \eta \circ \xi_1^{-1}|_{\xi(\mathcal{U} \cap \mathcal{V})}$$
$$\xi \circ \eta^{-1} = \xi_1 \circ \eta^{-1}|_{\eta(\mathcal{U} \cap \mathcal{V})}$$

are both C^∞-differentiable ∎

1.1.4 **Proposition.** *Let $\mathcal{T}_M$ be the collection of all the subsets of M which are unions of coordinates domains, together with the emptyset. $\mathcal{T}_M$ is a topology on M.*

Proof. Topology properties

$\emptyset\,,\, M \in \mathcal{T}_M$

$\{W_\alpha\} \subset \mathcal{T}_M \Longrightarrow \bigcup_\alpha W_\alpha \in \mathcal{T}_M$

are trivially satisfied. As for the last property,

$W_1\,,\, W_2 \in \mathcal{T}_M \Longrightarrow W_1 \cap W_2 \in \mathcal{T}_M\,,$

it is a direct consequence of the above lemma. ∎

Topology $\mathcal{T}_M$ will be called the *manifold topology* of M. It is a locally Euclidean topology, since, as we will prove, each coordinate domain is an open subset of M homeomorphic to an open subset of $\mathbb{R}^m$.

1.1.5 **Lemma.** *Given an admissible chart $(\mathcal{U}, \xi)$, for any open subset W of M contained in the domain $\mathcal{U}$, the restriction $\xi|_W$ is still an admissible chart.*

Proof. The proof follows the same pattern of the one of Lemma 1.1.3. ∎

1.1.6 **Proposition.** *Each admissible chart $(\mathcal{U}, \xi)$ defines a homeomorphism of its domain $\mathcal{U}$ onto its image $\xi(\mathcal{U})$.*

Proof. For any open subset $W \subset \mathcal{U}$, the restriction $\xi|_W$ belongs to $\mathcal{C}$ and then its image $\xi(W)$ is an open subset of $\xi(\mathcal{U})$. This proves that ξ is an open map. Let $\xi^{-1}(A) \subset \mathcal{U}$ be the inverse image of an open subset $A \subset \xi(\mathcal{U})$ of $\mathbb{R}^m$. The restriction $\xi|_{\xi^{-1}(A)}$ is an admissible chart on M since for any $(\mathcal{V}, \eta) \in \mathcal{C}$, the images

$$\eta\big(\mathcal{V} \cap \xi^{-1}(A)\big) = \eta\big(\mathcal{V} \cap \mathcal{U} \cap \xi^{-1}(A)\big) = \eta\big(\xi^{-1}\xi(\mathcal{V} \cap \mathcal{U}) \cap \xi^{-1}(A)\big) = \eta \circ \xi^{-1}\big(\xi(\mathcal{U} \cap \mathcal{V}) \cap A\big)$$

$$\xi\big(\mathcal{V}\big) \cap \xi^{-1}(A)\big) = \xi\big(\mathcal{V} \cap \mathcal{U} \cap \xi^{-1}(A)\big) = \xi(\mathcal{V} \cap \mathcal{U}) \cap A$$

are open subsets of $\mathbb{R}^m$. This proves that ξ is a continuous map too. ∎

It is worthwhile to remark that the locally Euclidean character of a manifold topology implies that a manifold M is *locally connected* (i.e., each point of M has a connected open neighbourhood). As a consequence, any connected component of M is an open subset of M. Moreover a manifold topology satisfies the first axiom of separation (i.e., any two points of M can be separated by two – not necessarily disjoint – open neighbourhoods) and the first axiom of countability (i.e., each point of M has a countable basis of open neighbourhoods).
In what follows, manifolds will be always meant to be Hausdorff and second-countable (1).

1.1.7 **Exercise.** *Any open subset W of an m–dimensional manifold M is an m–dimensional manifold. We will call W an open submanifold of M.*

---

(1) We recall that a Hausdorff manifold M is *locally compact*, i.e., for each point $x \in M$ and each open neighbourhood W of x, there exists an open neighbourhood V of x with compact clousure $\bar{V} \subset W$ (Cf. F.Brickell and R.S.Clark, Differentiable Manifolds, p.42).

1.2 Smooth mappings

Let M, N be smooth manifolds (with differential structures $\mathcal{C}_M$, $\mathcal{C}_N$ and dimensions m and n, respectively). Let

$$\Phi : M \longrightarrow N$$

be a mapping of M in N and $x \in M$. If $(\mathcal{U}, \xi) \in \mathcal{C}_M$ is a chart at x and $(\mathcal{V}, \eta) \in \mathcal{C}_N$ is a chart such that $\Phi(\mathcal{U}) \subset \mathcal{V}$, we define the *coordinate expression* of Φ around x by

$$\Phi_{\eta\xi} := \eta \circ \Phi \circ \xi^{-1} : \xi(\mathcal{U}) \longrightarrow \eta(\mathcal{V}) \ : \ \xi(y) \mapsto \eta\big(\Phi(y)\big).$$

The mapping Φ is said to be C^∞*–differentiable* at x, if there exists a coordinate expression $\Phi_{\eta\xi}$ around x which is C^∞ at $\xi(x)$. The above definition is, in fact, an intrinsic property of Φ, as follows from the following

1.2.1 **Proposition.** *If $\Phi_{\eta\xi}$ is C^∞ at $\xi(x)$, then any other coordinate expression $\Phi_{\eta'\xi'}$ around x is C^∞ at $\xi'(x)$.*

Proof. Notice that $\Phi_{\eta'\xi'} = \eta' \circ \Phi \circ \xi'^{-1}$ has a restriction to the open subset $\xi'(\mathcal{U}' \cap \mathcal{U}) \subset \xi'(\mathcal{U}')$ given by $(\eta' \circ \eta^{-1}) \circ \Phi_{\eta\xi} \circ (\xi \circ \xi'^{-1})$, whose C^∞–differentiability at $\xi'(x)$ follows from the C^∞–differentiability of transition functions $\xi \circ \xi'^{-1}$ and $\eta' \circ \eta^{-1}$. ∎

Now we show the local character of C^∞–differentiability.

1.2.2 **Proposition.**
(i) If Φ is C^∞ at x, so is its restriction to any open subset W containing x.
(ii) If a restriction of Φ to an open subset is C^∞ at x, so is Φ itself.

Proof. (i) Just notice that, if $\Phi_{\eta\xi}$ is a coordinate expression of Φ around x, then

$$(\Phi|_W)_{\eta\xi|_{\mathcal{U}\cap W}} = \Phi_{\eta\xi}|_{\xi(\mathcal{U}\cap W)}$$

is a coordinate expression of $\Phi|_W$ around x, since $\xi|_{\mathcal{U}\cap W}$ is an admissible chart on W.
(ii) The claim immediately follows from the fact that, if W is an open subset os M, any admissible chart on W is admissible on M too. ∎

1.2.3 **Exercises.**
(i) If Φ is C^∞ at x, it is continuous at x.
(ii) If Φ is C^∞ at x and Ψ is C^∞ at $\Phi(x)$, then $\Psi \circ \Phi$ is C^∞ at x.
(iii) Identity map id_M is C^∞ at any $x \in M$.
(iv) A function $f\colon A \subset \mathbb{R}^m \to \mathbb{R}^n$ (defined on an open subset A of $\mathbb{R}^m$) is C^∞ at x if, and only if, it is C^∞ at x in the Euclidean sense. ∎

A mapping $\Phi : M \to N$ is called a *smooth mapping* if it is C^∞ at every point of M (i.e., if it has C^∞–coordinate expressions in suitably many charts to cover M and $\Phi(M)$). Proposition 1.2.1 implies that all of the coordinate expressions of a smooth mapping Φ are C^∞. Proposition 1.2.2 extends to smooth mappings, and states
(i) the smoothness of the restriction of a smooth mapping Φ to any open subset of M;
(ii) the smoothness of a mapping Φ, if it admits smooth restrictions to suitably many open subsets to cover M.
Proposition 1.2.3 extends to smooth mappings in an obvious way.

A bijective, smooth mapping $\Phi\colon M \to N$ is called a *diffeomorphism*, if its inverse $\Phi^{-1}\colon N \to M$ is a smooth mapping. Diffeomorphic manifolds are structurally identical owing to the following

1.2.4 **Proposition.** *Let* $\Phi\colon M \to N$ *be a diffeomorphism. The map that takes each chart* ξ *with domain* $\mathcal{U} \subset M$ *to the chart* $\eta := \xi \circ \Phi^{-1}$ *with domain* $\Phi(\mathcal{U}) \subset N$*, defines a bijection between the differential structures* $\mathcal{C}_M$ *and* $\mathcal{C}_N$ *(which consequently have the same dimension).*

Proof. (i) First notice that, in the given map, to each chart $\xi \in \mathcal{C}_M$ there corresponds a chart $\eta \in \mathcal{C}_N$. This is due to the fact that, for any chart $(\mathcal{V}', \eta') \in \mathcal{C}_N$ such that $\mathcal{V}' \cap \Phi(\mathcal{U}) \neq \emptyset$, the transition functions

$$\eta' \circ \eta^{-1}|_{\eta(\Phi(\mathcal{U})\cap\mathcal{V}')} = \eta' \circ \Phi \circ \xi^{-1}|_{\xi(\mathcal{U}\cap\Phi^{-1}(\mathcal{V}'))}$$
$$\eta \circ \eta'^{-1}|_{\eta'(\mathcal{V}'\cap\Phi(\mathcal{U}))} = \xi \circ \Phi^{-1} \circ \eta'^{-1}|_{\eta'(\mathcal{V}'\cap\Phi(\mathcal{U}))}$$

are C^∞, being the coordinate expressions of Φ and Φ^{-1} in admissible charts on M and N.
(ii) Then notice that, in the given map, any chart $\eta \in \mathcal{C}_N$ corresponds to a unique chart $\xi = \eta \circ \Phi$ and, through the same reasoning as in (i), one can check that $\xi \in \mathcal{C}_M$. ∎

Any bijection onto a manifold can be turned into a diffeomorphism as follows.

1.2.5 **Proposition.** *Let* $\Phi\colon M \to N$ *be a bijection of a smooth manifold* M *onto a set* N*. There exists a unique differential structure on* N *such that* Φ *is a diffeomorphism.*

Proof. The differential structure $\mathcal{C}_N$ searched, is obtained from $\mathcal{C}_M$ through the map given in Proposition 1.2.4. ∎

A nice example of diffeomorphism is given in the following

1.2.6 **Proposition.**
(i) Any admissible chart $(\mathcal{U}, \xi)$ *on* M *defines a diffeomorphism between the open submanifolds* $\mathcal{U} \subset M$ *and* $\xi(\mathcal{U}) \subset \mathbb{R}^m$.
(ii) Conversely, any diffeomorphism ξ *between the open submanifolds* $\mathcal{U} \subset M$ *and* $\xi(\mathcal{U}) \subset \mathbb{R}^m$ *defines an admissible chart on* M.

Proof. (i) It is enough to remark that the coordinate expressions of ξ and ξ^{-1} in charts ξ on $\mathcal{U}$ and $\mathrm{id}_{\xi(\mathcal{U})}$ on $\xi(\mathcal{U})$, are C^∞ (for they both reduce to $\mathrm{id}_{\xi(\mathcal{U})}$).
(ii) It is enough to remark that, for any admissible chart ξ' whose domain $\mathcal{U}'$ encounters $\mathcal{U}$, the transition functions $\xi' \circ \xi^{-1}$ and $\xi \circ \xi'^{-1}$, being composition of smooth mappings, are smooth mappings. ∎

1.3 Bump functions

Let M be a smooth manifold and $C^\infty(M)$ the algebra of real-valued smooth functions on M. The existence of a special kind of functions in $C^\infty(M)$ is clamed (2) in the following statement: for any point $x \in M$ and any open neighbourhood W of x, there exixts a function $\beta \in C^\infty(M)$ which takes the constant value 1 on an open neighbourhood of x and has its support (3) contained in W. Te function β is called a *bump function* at x with support in W. The above statement implies

$$\begin{cases} \beta|_{\mathcal{V}} = 1, & \mathcal{V} \subset W \\ \beta|_{M-W} = 0. \end{cases}$$

(2) See B.O'Neill, Semi-Riemannian Geometry, p.6.
(3) $\operatorname{supp}\beta :=$ closure of $\{x \in M \mid \beta(x) \neq 0\}$.

Bump functions are used to 'extend' smooth mappings defined on an open neighbourhood of a point, to a smooth mapping on M . For instance, let $\varphi \in C^\infty(W)$. Put

$$f(y) = \begin{cases} \beta(y)\varphi(y), & \text{if } y \in W \\ 0, & \text{if } y \in M - W. \end{cases}$$

We have

$$f \in C^\infty(M)$$

(since $f|_W = \beta|_W\varphi$ and $f|_{M-\operatorname{supp}\beta} = 0$ are smooth restrictions of f to open subsets W and $M - \operatorname{supp}\beta$, which cover M), and

$$f|_{\mathcal{V}} = \varphi|_{\mathcal{V}}\,.$$

2 Tangent spaces and tangent mappings

2.1 Tangent spaces

Let M be an m–dimensional smooth manifold and $x \in M$. Any $\mathbb{R}$-linear map

$$v\colon C^\infty(M) \longrightarrow \mathbb{R}$$

which obeys the Leibniz rule at x

$$v(fg) = v(f)g(x) + f(x)v(g) \qquad \forall f,\, g \in C^\infty(M)$$

is called a *derivation* of $C^\infty(M)$ at x. The local character of a derivation v is pointed out in the following proposition.

2.1.1 **Proposition.** *Let $h \in C^\infty(M)$. If $h|_{\mathcal{U}} = 0$ in some open neighbourhood $\mathcal{U}$ of x, then $v(h) = 0$.*

Proof. Let β be a bump function at x with support in $\mathcal{U}$. We have $h = (1-\beta)h$ both on $\mathcal{U}$ (where h vanishes) and on $M - \mathcal{U}$ (where β vanishes). The Leibniz rule then implies

$$v(h) = v(1-\beta)h(x) + (1-\beta(x))v(h) = 0$$

since $h(x) = 0$ and $\beta(x) = 1$, and $\mathbb{R}$-linearity implies the thesis. ■

It follows that a derivation v at x induces a derivation on any open neighbourhood W of x, by putting, for each $\varphi \in C^\infty(W)$,

$$v(\varphi) := v(f)$$

with any $f \in C^\infty(M)$ equal to φ around x.
Lastly we remark that

2.1.2 **Corollary.** *If f is a constant function, then $v(f) = 0$.*

Proof. If $c \in \mathbb{R}$ denotes the constant value of f and $1 \in C^\infty(M)$ the unit function on M, from $f = c1$ it follows that $v(f) = v(c1) = cv(1)$. But $v(1) = v(1 \cdot 1) = v(1) + v(1)$ that is $v(1) = 0$, whence the statement. ■

Let T_xM be the set of all the derivations of $C^\infty(M)$ at x. It is easy to prove the following

2.1.3 **Proposition.** *Let $u, v \in T_xM$ and $a \in\mathbb{R}$ and define, for any $f \in C^\infty(M)$,*

$$(u+v)f = u(f) + v(f)$$
$$(av)f = av(f)$$

With these operations, T_xM is given a structure of real vector space. ∎

The space T_xM, endowed with the above structure, is called the *tangent space* of M at x (and any $v \in T_xM$, a *tangent vector* of M at x). As to the dimension of T_xM, we will show that

$$\dim T_xM = m = \dim M.$$

To this end, we consider an admissible chart $\xi = (x^i)_{i=1,\dots,m}$ at x and define

$$\left.\frac{\partial}{\partial x^i}\right|_x : C^\infty(M) \longrightarrow \mathbb{R} \qquad (i = 1,\dots,m)$$

by putting, for each $f \in C^\infty(M)$,

$$\left.\frac{\partial}{\partial x^i}\right|_x (f) := \left.\frac{\partial f}{\partial x^i}\right|_x := \left.\frac{\partial(f\circ\xi^{-1})}{\partial x^i}\right|_{\xi(x)}$$

(*partial derivative* at x with respect to the i-th coordinate function x^i). From $\mathbb{R}$-linearity and Leibniz rule of partial derivatives in Euclidean calculus, it follows that

$$\left.\frac{\partial}{\partial x^i}\right|_x \in T_xM.$$

The *coordinate vectors* $\left\{ \frac{\partial}{\partial x^i}\big|_x \right\}$ associated with ξ, are a basis of T_xM, owing to the following

2.1.4 **Theorem.** *Any vector $v \in T_xM$ can be uniquely written as a linear combination*

$$v = v^i \left.\frac{\partial}{\partial x^i}\right|_x$$

with components

$$v^i := v(x^i)$$

Proof. (i) We start with some preliminary calculations concerning C^∞ real functions on $\mathbb{R}^m$. Let B_0^r be the open ball with center in the origin $0 \in\mathbb{R}^m$ and radius $r > 0$. Let $\bar{x}$ be a point of B_0^r. If $\bar{x} \neq 0$ (and obviously $|\bar{x}| < r$), consider the open interval $I = (-r/|\bar{x}|\, , r/|\bar{x}|)$ containing $[0,1]$ and define in I the C^∞ function $\gamma_{\bar{x}}(t) = t\bar{x}$. If $\bar{x} = 0$, the same definition yields the null function. In both cases, the image of $\gamma_{\bar{x}}$ is contained in B_0^r and $\gamma_{\bar{x}}(0) = 0$, $\gamma_{\bar{x}}(1) = \bar{x}$. The projections $\gamma_{\bar{x}}^i := pr^i \circ \gamma_{\bar{x}} = t\bar{x}^i$, $(i = 1,\dots,m)$ have constant derivatives

$$\frac{d}{dt}\gamma_{\bar{x}}^i = \bar{x}^i.$$

Now consider a C^∞ real function $F\colon B_0^r \to\mathbb{R}$. Composition $F\circ\gamma_{\bar{x}} = F(\gamma_{\bar{x}}^1,\dots,\gamma_{\bar{x}}^m)$ is a well defined C^∞ function with derivative

$$\frac{d}{dt}(F\circ\gamma_{\bar{x}}) = \left(\frac{\partial F}{\partial x^i}\circ\gamma_{\bar{x}}\right)\bar{x}^i.$$

From elementary integral calculus, then we have

$$F(\bar{x}) - F(0) = \int_0^1 \left(\frac{d}{dt} F \circ \gamma_{\bar{x}} \right) dt = \left(\int_0^1 \left(\frac{\partial F}{\partial x^i} \circ \gamma_{\bar{x}} \right) dt \right) \bar{x}^i = F_i(\bar{x}) pr^i(\bar{x})$$

where

$$F_i(\bar{x}) := \int_0^1 \left(\frac{\partial F}{\partial x^i} \circ \gamma_{\bar{x}} \right) dt$$

is a C^∞ function on B_0^r. So, on B_0^r, we have

$$F = F(0) + F_i pr^i.$$

(ii) A similar result will be now obtained around $x \in M$ for any $f \in C^\infty(M)$. Let $(\mathcal{U}, \xi)$ be an admissible chart at x, with $\xi(x) = 0$ and $\xi(\mathcal{U}) = B_0^r$ [4]. Then let $F := f \circ \xi^{-1}$ be the coordinate expression of f in ξ. From (i) it follows that

$$f|_{\mathcal{U}} = F \circ \xi = \big(F(0) + F_i pr^i\big) \circ \xi = F(0) + (F_i \circ \xi) x^i.$$

If we recall that $F(0) = f(x)$ and we put $f_i := F_i \circ \xi \in C^\infty(\mathcal{U})$, we have

$$f|_{\mathcal{U}} = f(x) + f_i x^i.$$

(iii) Now, for any $v \in T_x M$ and $f \in C^\infty(M)$,

$$\begin{aligned} v(f) &= v(f|_{\mathcal{U}}) = v\big(f(x) + f_i x^i\big) = v\big(f(x)\big) + v(f_i x^i) = v(f_i x^i) = v(f_i) x^i(x) + f_i(x) v(x^i) \\ &= f_i(x) v(x^i) \end{aligned}$$

since $x^i(x) = 0$. In particular

$$\left.\frac{\partial}{\partial x^i}\right|_x (f) = f_j(x) \left.\frac{\partial}{\partial x^i}\right|_x (x^j) = f_j(x) \delta_i^j = f_i(x),$$

then

$$v(f) = v(x^i) \left.\frac{\partial}{\partial x^i}\right|_x (f),$$

whence the stated decomposition of v.

(iv) Lastly we remark that the uniqueness of the above decomposition is due to the linear independence of coordinate vectors, which is easily shown by

$$a^i \left.\frac{\partial}{\partial x^i}\right|_x = 0 \Longrightarrow a^i \left.\frac{\partial}{\partial x^i}\right|_x (x^j) = 0 \Longrightarrow a^j = 0$$

for all $j = 1, \ldots, m$. ■

By applying the decomposition law to coordinate vectors, we have

[4] Such a *spherical chart* can be obtained from any admissible chart at x through a translation and a restriction.

2.1.5 **Corollary.**
(i) The bases $\left\{ \frac{\partial}{\partial x^i}\Big|_x \right\}$ *and* $\left\{ \frac{\partial}{\partial y^j}\Big|_x \right\}$ *associated with the admissible charts* $\xi = (x^i)$ *and* $\eta = (y^j)$ *are related to each other by the chain rule*

$$\frac{\partial}{\partial y^j}\Big|_x = \frac{\partial x^i}{\partial y^j}\Big|_x \frac{\partial}{\partial x^i}\Big|_x$$

(ii) Consequently, the controvariant transformation law of the components of a vector v *is given by*

$$v(x^i) = \frac{\partial x^i}{\partial y^j}\Big|_x v(y^j)$$

■

2.1.6 **Note.** We will denote by

$$\xi'_x : T_xM \longrightarrow \mathbb{R}^m \; : \; v = v^i \frac{\partial}{\partial x^i}\Big|_x \mapsto \bar{v} = (v^i)$$

the isomorphism defined by the coordinate basis $\left\{ \frac{\partial}{\partial x^i}\Big|_x \right\}$ at x. ■

2.1.7 **Remark.** Let W be an open submanifold of M and $x \in W$.
(i) If, for any $u \in T_xW$, we define $v: C^\infty(M) \to \mathbb{R}$ by putting $v(f) := u(f|_W)$, we have $v \in T_xM$.
(ii) The consequent map $u \in T_xW \mapsto v \in T_xM$ is a (canonical) isomorphism. We shall usually put $T_xW = T_xM$. ■

2.2 Tangent mappings

Let $\Phi: M \to N$ be a smooth mapping and $x \in M$. For any $v \in T_xM$, define

$$T_x\Phi \cdot v \; : C^\infty(N) \longrightarrow \mathbb{R}$$

by putting

$$(T_x\Phi \cdot v)(f) = v(f \circ \Phi).$$

2.2.1 **Lemma.** $T_x\Phi \cdot v \in T_{\Phi(x)}N$.

Proof. Let us check, for instance, that $w := T_x\Phi \cdot v$ obeys the Leibniz rule at $y := \Phi(x)$. To this purpose, let $f, g \in C^\infty(N)$. We have

$$w(fg) = v(fg \circ \Phi) = v(f \circ \Phi)g(\Phi(x)) + f(\Phi(x))v(g \circ \Phi) = w(f)g(y) + f(y)w(y)$$

■

The consequent mapping

$$T_x\Phi \; : T_xM \longrightarrow T_{\Phi(x)}N$$

is called the *tangent mapping* of Φ at x.

2.2.2 Proposition.
(i) $T_x\Phi$ *is a linear mapping.*
(ii) $T_x\mathrm{id}_M = \mathrm{id}_{T_xM}$.
(iii) (Chain rule) $T_x(\Psi \circ \Phi) = T_{\Phi(x)}\Psi \circ T_x\Phi$.

Proof. Let us prove, for instance, property (iii). Let $\Phi\colon M \to N$, $\Psi\colon N \to P$ be smooth mappings. For each $v \in T_xM$ and each $f \in C^\infty(P)$, we have

$$\big(T_x(\Psi \circ \Phi) \cdot v\big)f = v(f \circ \Psi \circ \Phi) = (T_x\Phi \cdot v)(f \circ \Psi) = (T_{\Phi(x)}\Psi \circ T_x\Phi \cdot v)f$$ ∎

2.2.3 Corollary. *If* $\Phi\colon M \to N$ *is a diffeomorphism, then* $T_x\Phi$ *is an isomorphism, whose inverse is given by*

$$(T_x\Phi)^{-1} = T_{\Phi(x)}\Phi^{-1}$$

Proof. The statement follows from the Proposition 2.2.2(ii),(iii), applied to $\mathrm{id}_M = \Phi^{-1} \circ \Phi$ and $\mathrm{id}_N = \Phi \circ \Phi^{-1}$. ∎

2.2.4 Coordinate expression.
Let $\Phi\colon M \to N$ be a smooth mapping. Let $\xi = (x^i)_{i=1,\dots,m=\dim M}$ and $\eta = (y^\alpha)_{\alpha=1,\dots,n=\dim N}$ be admissible charts at x and $\Phi(x)$, respectively, where Φ can be given a coordinate expression $\Phi_{\eta\xi}$. For any

$$v = v^i \left.\frac{\partial}{\partial x^i}\right|_x \in T_xM$$

put

$$w := T_x\Phi \cdot v = w^\alpha \left.\frac{\partial}{\partial y^\alpha}\right|_{\Phi(x)} \in T_{\Phi(x)}N.$$

By linearity, we have

$$w = v^i \left(T_x\Phi \cdot \left.\frac{\partial}{\partial x^i}\right|_x\right) = v^i \left(T_x\Phi \cdot \left.\frac{\partial}{\partial x^i}\right|_x\right)(y^\alpha) \left.\frac{\partial}{\partial y^\alpha}\right|_{\Phi(x)} = v^i \left.\frac{\partial(y^\alpha \circ \Phi)}{\partial x^i}\right|_x \left.\frac{\partial}{\partial y^\alpha}\right|_{\Phi(x)},$$

and then

$$w^\alpha = v^i \left.\frac{\partial(y^\alpha \circ \Phi)}{\partial x^i}\right|_x = v^i \left.\frac{\partial(y^\alpha \circ \Phi \circ \xi^{-1})}{\partial x^i}\right|_{\xi(x)} = v^i \left.\frac{\partial \Phi^\alpha_{\eta\xi}}{\partial x^i}\right|_{\xi(x)}$$

(with $\Phi^\alpha_{\eta\xi} := pr^\alpha \circ \Phi_{\eta\xi}$), or equivalently

$$\bar{w} = d_{\xi(x)}\Phi_{\eta\xi} \cdot \bar{v}$$

(with $\bar{w} := (w^\alpha)$ and $\bar{v} := (v^i)$). So we have the following commutative diagram

$$\begin{array}{ccc} v \in T_xM & \xrightarrow{\;T_x\Phi\;} & w \in T_{\Phi(x)}N \\ \Big\downarrow{\scriptstyle \xi'_x} & & \Big\downarrow{\scriptstyle \eta'_{\Phi(x)}} \\ \bar{v} \in \mathbb{R}^m & \xrightarrow[\;d_{\xi(x)}\Phi_{\eta\xi}\;]{} & \bar{w} \in \mathbb{R}^n \end{array}$$

(where ξ'_x and $\eta'_{\Phi(x)}$ denote the isomorphisms defined by the coordinate bases). We will say that $d_{\xi(x)}\Phi_{\eta\xi}$ is the coordinate expression of $T_x\Phi$ in ξ, η.
Notice that, for a smooth function $f : A \subset \mathbb{R}^m \to \mathbb{R}^n$, the Euclidean differential d_xf is the coordinate expression of the tangent map T_xf in charts id_A and $\mathrm{id}_{\mathbb{R}^n}$. ∎

2.2.5 Remarks.
(i) Let $f \in C^\infty(M)$. For any $x \in M$, the composition of the tangent map $T_x f : T_x M \to T_{f(x)}\mathbb{R}$ with the natural isomorphism $\mathrm{id}'_{f(x)} : T_{f(x)}\mathbb{R} \to \mathbb{R}$ (defined by the chart $\mathrm{id}_\mathbb{R}$ on $\mathbb{R}$) yields the *differential*

$$d_x f := \mathrm{id}'_{f(x)} \circ T_x f \; : T_x M \to \mathbb{R}$$

which will usually replace $T_x f$. Notice that $d_x f$ is a linear form on the vector space $T_x M$, i.e., an element of the dual space $T_x^* M$, whose action on any $v \in T_x M$ is $d_x f \cdot v = \mathrm{id}'_{f(x)}(T_x f \cdot v)$, that is

$$d_x f \cdot v = v(f).$$

(ii) Now consider a smooth mapping $F = (F^1, \ldots, F^n) : M \longrightarrow \mathbb{R}^n$. As above, for any $x \in M$, define the differential

$$d_x F := \mathrm{id}'_{F(x)} \circ T_x F \; : T_x M \to \mathbb{R}^n$$

(with $\mathrm{id}_{\mathbb{R}^n} = (y^\alpha)_{\alpha=1,\ldots,n}$). For any $v \in T_x M$,

$$\begin{aligned} d_x F \cdot v &= \mathrm{id}'_{F(x)}\big(T_x F \cdot v\big) = \big(T_x F \cdot v\,(y^\alpha)\big) = \big(v(y^\alpha \circ F)\big) \\ &= \big(v(F^\alpha)\big) = \big(d_x F^\alpha \cdot v\big) = \big(d_x F^1, \ldots, d_x F^n\big) \cdot v \end{aligned}$$

that is

$$d_x F = (d_x F^1, \ldots, d_x F^n). \qquad \blacksquare$$

Let M and N be smooth manifolds of dimensions m and n, respectively, and $\Phi : M \to N$ a smooth mapping. We define the *rank* of Φ at $x \in M$ as

$$\mathrm{rank}\ \Phi(x) := \mathrm{rank}\ T_x\Phi = \dim \mathrm{Im}\ T_x\Phi.$$

According to the well known rank theorem for linear mappings, it is

$$\dim T_x M = \dim \ker T_x\Phi + \dim \mathrm{Im}\ T_x\Phi.$$

As a consequence,

$$\mathrm{rank}\ \Phi(x) \le m, n.$$

Also notice that, for any coordinate expression $\Phi_{\eta\xi}$,

$$\mathrm{rank}\ \Phi(x) = \mathrm{rank}\ \Phi_{\eta\xi}(\xi(x)) = \mathrm{rank}\ d_{\xi(x)}\Phi_{\eta\xi}$$

since $T_x\Phi$ and $T_{\xi(x)}\Phi_{\eta\xi}$ are both related to $d_{\xi(x)}\Phi_{\eta\xi}$ by isomorphisms [(7)].
The main result of rank theory is the existence of *adapted charts*, as stated in the following theorem.

2.2.6 Rank Theorem. *Let rank $\Phi(x) = k$. Then there exist admissible charts*

$$\xi = (x^1, \ldots, x^m) : \mathcal{U} \to \mathbb{R}^m$$
$$\eta = (y^1, \ldots, y^n) : \mathcal{V} \to \mathbb{R}^n$$

at x and $\Phi(x)$, respectively, such that
(o) $\Phi(\mathcal{U}) \subset \mathcal{V}$
(i) $y^i \circ \Phi|_\mathcal{U} = x^i \qquad \forall i = 1, \ldots, k.$
If the rank of Φ is k in some neighbourhood of x (and $k < n$), we can further arrange that
(ii) $y^j \circ \Phi|_\mathcal{U} = const. \qquad \forall j = k+1, \ldots, n\,.$
Conversely, the existence of charts as above $(o) - (ii)$, implies that the rank of Φ is k in some open neighbourhood of x.

Proof. Cf. M.Boothby *An introduction to differentiable manifolds and Riemmanian geometry.* ■

(7) See Coordinate expression 2.2.4.

2.3 Smooth curves

A smooth mapping $c\colon I \to M$, of an open interval $I \subset\mathbb{R}$ in a smooth manifold M, is said to be a *smooth curve*, or *motion*, in M (and $c(I)$ its *orbit*). Let $\frac{d}{dt}\big|_t \in T_t I = T_t\mathbb{R}$ be the coordinate vector associated with chart $id_{\mathbb{R}}$ on $\mathbb{R}$. The *time derivative* of c at $t \in I$, defined by

$$\dot{c}(t) := T_t c \cdot \frac{d}{dt}\bigg|_t \in T_{c(t)}M \,,$$

is said to be the *tangent vector* of c at the point $c(t)$, or the *velocity* of c at time t.

2.3.1 **Coordinate expression.**
Let $t \in I$ and let ξ be an admissible chart at $c(t)$. By continuity, there exists an open interval containing t and contained in I, whose c-image lies in the domain of ξ. There c has a coordinate expression $(x^i \circ c)_{i=1,\ldots,m=\dim M}$. As to $\dot{c}(t)$, it can be expressed as

$$\dot{c}(t) = \big(\dot{c}(t)(x^i)\big)\, \frac{\partial}{\partial x^i}\bigg|_{c(t)}$$

with components

$$\dot{c}(t)(x^i) = \left(T_t c \cdot \frac{d}{dt}\bigg|_t\right)(x^i) = \frac{d}{dt}\bigg|_t (x^i \circ c) = \frac{dc^i)}{du}\bigg|_t =: \dot{c}^i(t),$$

So we have

$$\dot{c}(t) = \dot{c}^i(t)\, \frac{\partial}{\partial x^i}\bigg|_{c(t)} .$$

■

2.3.2 **Remarks.**
(i) Notice that, for $c : I \to\mathbb{R}$, we have

$$\dot{c}(t) = \frac{dc}{dt}\bigg|_t \frac{d}{dt}\bigg|_{c(t)}$$

i.e., the Euclidean derivative $\frac{dc}{dt}|_t$ is the natural component of the tangent vector $\dot{c}(t)$.
(ii) Similarly, for $c : I \to\mathbb{R}^m$, we have

$$\dot{c}(t) = \frac{d(x^i \circ c)}{dt}\bigg|_t \frac{\partial}{\partial x^i}\bigg|_{c(t)}$$

i.e., the Euclidean derivative $\frac{dc}{dt}\big|_t := \left(\frac{d(x^i \circ c)}{dt}\Big|_t\right)$ is the m-tuple of the natural components of the tangent vector $\dot{c}(t)$. ■

2.3.3 **Proposition.** *For any $x \in M$ and $v \in T_xM$, there exists a smooth curve $c: I \to M$ with $0 \in I$, such that $\dot{c}(0) = v$.*

Proof. Let $\xi = (x^i)$ be an admissible chart at x, with domain $\mathcal{U}$. Consider the rectilinear motion in $\mathbb{R}^m$

$$\gamma(t) = \xi(x) + t\xi'_x(v).$$

Since $\gamma(0) = \xi(x)$, by continuity there exists an open interval I whose γ-image lies in $\xi(\mathcal{U})$; so, through restriction of γ to I, one obtains a smooth mapping $\gamma|_I : I \to \xi(\mathcal{U})$. Now put

$$c := \xi^{-1} \circ \gamma|_I : I \longrightarrow M.$$

The mapping c is a smooth curve in M with the required property. This follows from the coordinate expression

$$c^i = x^i \circ c = (pr^i \circ \xi) \circ (\xi^{-1} \circ \gamma|_I) = pr^i \circ \gamma|_I$$

and then

$$c^i(t) = x^i(x) + tv(x^i)$$

which yields $\dot{c}^i(t) = v(x^i)$. So we have

$$\dot{c}(0) = \dot{c}^i(0) \left.\frac{\partial}{\partial x^i}\right|_{c(0)} = v(x^i) \left.\frac{\partial}{\partial x^i}\right|_x = v$$ ∎

2.3.4 **Remark.** As a consequence of Proposition 2.3.3, we have that, for any $v \in T_xM$ and $f \in C^\infty(M)$,

$$v(f) = \dot{c}(0)(f) = \left(T_0 c \cdot \left.\frac{d}{dt}\right|_0\right)(f) = \left.\frac{d}{dt}\right|_0 (f \circ c) = \left.\frac{d(f \circ c)}{dt}\right|_0$$

i.e., $v(f)$ is the directional derivative of f at x along any smooth curve c to which v is tangent. ∎

2.3.5 **Remark.** Let $k = \Phi \circ c : I \to N$ whit $\Phi : M \to N$. Then, by applying to the chain rule of tangent mappings, for any $t \in I$, we have

$$\dot{k}(t) = T_{c(t)}\Phi \cdot \dot{c}(t).$$

As a consequence of Propositions 2.3.3 then, notice that one can always reduce the calculus of a tangent mapping $w = T_x\Phi \cdot v$, to a time derivative $w = \dot{k}(0)$ (where $k = \Phi \circ c$ and $\dot{c}(0) = v$). ∎

---

3 Immersions and submersions

3.1 Submanifolds

Let S, M be smooth manifolds of dimensions s, m respectively, with $s \leq m$ and

$$\jmath : S \longrightarrow M$$

a smooth mapping. If the tangent map of $\jmath$ at a point $x \in S$ is injective, $\jmath$ is said to be an *immersion* at x. In such a case, from $\dim \ker T_x\jmath = 0$, it follows that $\operatorname{rank} \jmath(x) = s = \max$, and then Rank Theorem 2.2.6 reads as follows.

3.1.1 **Proposition.** *If $\jmath$ is an immersion at x, we can find admissible charts $\xi : \mathcal{U} \to \mathbb{R}^s$, $\eta : \mathcal{V} \to \mathbb{R}^m$ at x, $\jmath(x)$ such that*
(o) $\jmath(\mathcal{U}) \subset \mathcal{V}$
(i) $\eta \circ \jmath|_{\mathcal{U}} = (\xi\,,\, c)$ *where $c : \mathcal{U} \to \mathbb{R}^{m-s}$ denotes a smooth mapping.*
If $\jmath$ is an immersion in some open neighbourhood of x (and $s < m$), we can further arrange that
(ii) $c = const.$
Conversely, the existence of charts as above $(o) - (ii)$, implies that $\jmath$ is an immersion in some open neighbourhood of x. ∎

Any differential structure on a subset S of M which makes the inclusion mapping $\jmath : S \hookrightarrow M$ an immersion (i.e., an immersion at each point of S) is called a *submanifold structure* on S, and S, endowed with such a structure, an (*immersed*) *submanifold* of M.

3.1.2 **Remarks.**
(i) Any injective immersion $\imath : S \to M$ gives rise to a submanifold of M, namely $\imath(S)$ with the differential structure which makes the induced mapping $\tilde{\imath} : S \to \imath(S)$ a diffeomorphism (8).
(ii) If S is a submanifold of M and M is a submanifold of N, then S is a submanifold of N. In fact, is we denote by $\jmath : S \hookrightarrow M$ and $\jmath' : M \hookrightarrow N$ the inclusion maps, then, at each point of S, rank $(\jmath' \circ \jmath) = \dim S$. ∎

First we will study the topology of a submanifold. As a preliminary, we recall the following well known result on continuity.
Let $\Phi : P \to M$ be a continuous mapping, and consider any subset S of M, such that $\Phi(P) \subset S$. With respect to the subspace topology on S,the induced mapping $\tilde{\Phi} : P \to S$ is continuous. (9).

3.1.3 **Proposition.** *The manifold topology of a submanifold is finer then its subspace topology.*

Proof. Let S be a submanifold of M. Then S is endowed both with the manifold topology, that we will denote by S_τ, and with the subspace topology, that we will denote by S_σ. The continuity of the immersion $\jmath : S_\tau \hookrightarrow M$ implies the continuity of the induced mapping $\tilde{\jmath} = id_S : S_\tau \longrightarrow S_\sigma$, and then the claim. ∎

Now we study the behaviour of submanifolds with respect to smooth mappings.
If $\Phi : P \to M$ is a smooth mapping, then its restriction $\Phi|_S$ to any submanifold S is still a smooth mapping (since $\Phi|_S = \Phi \circ \jmath$). On the contrary, the induced mapping $\tilde{\Phi} : P \to S$ with values in a submanifold S such that $\Phi(P) \subset S$ will not generally be smooth – not even continuous, continuity being guaranteed by the subspace topology of S but not generally by the finer manifold topology of S (10). If, for each smooth mapping $\Phi : P \to M$ with values in a given submanifold S of M, the induced mapping $\tilde{\Phi} : P \to S$ is smooth, then we will call S a *smoothness preserving* submanifold. For such a submanifold, the following remarkable property holds true.

3.1.4 **Proposition.** *On any subset of a manifold, there exists at most one smoothness preserving submanifold structure.*

Proof. Let $S_1 := (S, \mathcal{C}_1)$ and $S_2 := (S, \mathcal{C}_2)$ be two smoothness preserving submanifold structures on S and consider the immersions $\jmath_1 : S_1 \hookrightarrow M$ and $\jmath_2 : S_2 \hookrightarrow M$, respectively. On the one hand,

---

(8) See Proposition 1.2.5.
(9) If $\mathcal{V} = W \cap S$ with W open subset of M, then $\tilde{\Phi}^{-1}(\mathcal{V}) = \tilde{\Phi}^{-1}(W \cap S) = \Phi^{-1}(W)$ is an open subset of P.
(10) For a counterexample, see F.Brickell and R.S.Clark, op. cit., p.76.

we recall that $\jmath_1$ (resp. $\jmath_2$) is a smooth mapping with values in S_2 (risp. S_1) and then the induced mapping $\tilde{\jmath}_1 : S_1 \to S_2$ (resp. $\tilde{\jmath}_2 : S_2 \to S_1$) is smooth as well. On the other hand, from the set theoretical point of view we have $\tilde{\jmath}_1 = id_S = \tilde{\jmath}_2$. We conclude that $\tilde{\jmath}_1$ is a diffeomorphism whose inverse is $\tilde{\jmath}_2$ and then $\mathcal{C}_1 = \mathcal{C}_2$. ■

Let us now focus on the main type of smoothness preserving submanifold, defined as follows. An immersion $\jmath : S \to M$ such that the induced mapping $\tilde{\jmath} : S \to \jmath(S)$ is a homomorphism onto the topological subspace $\jmath(S) \subset M$ is called an *embeddeing*. A subset S of M, carrying a differential structure which makes inclusion map $\jmath : S \hookrightarrow M$ an embedding is called an *embedded* (or *regular*) submanifold.

3.1.5 **Remarks.**
(i) Any embedding $\imath : S \to M$ gives rise to an embedded submanifold of M, namely $\jmath(S)$ with the differential structure given in Remark 3.1.2(i).
(ii) Remark 3.1.2(ii) trivially extends to embedded submanifolds. ■

The peculiarity of an embedded submanifold clearly lies in its manifold topology, as is comfirmed by the following

3.1.6 **Proposition.** *A submanifold S is embedded iff its manifold topology coincides with its subspace topology.* ■

3.1.7 **Open submanifolds.** Let W (with differential structure $\mathcal{O}_W$) be an open submanifold of M [11]. For each $x \in W$, we can find an m–chart $(\mathcal{U}, \xi)$ with $x \in \mathcal{U} \subset W$, which is admissible on both W and M. This implies that $\jmath : W \hookrightarrow M$ is an immersion [12]. Moreover the manifold topology of W coincides with its subspace topology. So W is an embedded submanifold and $\dim W = \dim M$. Notice that the open submanifolds are the only submanifolds of M whose dimension is $\dim M$. In order to check this, let W be a submanifold such that $\dim W = \dim M = m$ and let $\jmath : W \hookrightarrow M$ be its immersion into M. For each $x \in W$, we can find admissible m–charts $(\mathcal{U}, \xi)$ on W and $(\mathcal{V}, \eta)$ on M with $x \in \mathcal{U} \subset \mathcal{V}$ and $\eta \circ \jmath|_{\mathcal{U}} = \xi$. From $\eta(\mathcal{U}) = \eta\big(\jmath(\mathcal{U})\big) = \xi(\mathcal{U})$ or equivalently $\mathcal{U} = \eta^{-1}\big(\xi(\mathcal{U})\big)$, we draw that $\mathcal{U}$ is an open subset of M. So W, being union of open subsets such as $\mathcal{U}$, is an open subset of M. Also notice that, owing to the above equality of dimensions, for each point x of an open submanifold W the tangent mapping $T_x\jmath \in T_xW \to T_xM$ is an isomorphism, namely the canonical isomorphism [13], since, for any $f \in C^\infty(M)$, $(T_x\jmath \cdot u)f = u(f \circ \jmath) = u(f|_W)$. ■

3.1.8 **Compact submanifolds.** Let C be a compact submanifold of M (i.e., compact in its own manifold topology) and $\jmath : C \hookrightarrow M$ its immersion into M. The image $\jmath(C)$, as a topological subspace of a manifold M (which is meant to be Hausdorff), is a Hausdorff space. The induced mapping $\tilde{\jmath} : C \to \jmath(C)$ is then a continuous bijection of a compact space onto a Hausdorff space, and therefore a homeomorphism [14]. So C is an embedded submanifold. ■

As already announced, we have the following

---

(11) See Exercise 1.1.7
(12) See Proposition 3.1.1.
(13) See Remark 2.1.7.
(14) A continuous bijection $h : X \to Y$ of a compact space X onto a Hausdorff space Y, is an open mapping. Indeed, given an open subset A of X, the closed subset $X - A$ of the compact space X is compact itself and so is its continuous image $h(X - A) = h(X) - h(A) = Y - h(A)$; a compact subset, such as $Y - h(A)$, of a Hausdorff space Y is closed, and then $h(A)$ is open.

3.1.9 **Proposition.** *An embedded submanifold is smoothness preserving.*

Proof. Let $\Phi : N \to M$ be a smooth mapping with values in an embedded submanifold S and $\tilde{\Phi} : N \to S$ the induced mapping. We will prove that $\tilde{\Phi}$ is smooth.
(i) As $\jmath : S \hookrightarrow M$ is an immersion, for each $y \in N$ we can find charts $(\mathcal{U}, \xi)$ and $(\mathcal{V}, \eta)$ at $\tilde{\Phi}(y)$ and $\jmath\big(\tilde{\Phi}(y)\big)$ adapted to $\jmath$, i.e., $\mathcal{U} \subset \mathcal{V}$ and $\eta \circ \jmath|_{\mathcal{U}} = (\xi, c)$ [15].
(ii) As the manifold topology of S coincides with its subspace topology, $\tilde{\Phi}$ is continuous [16]. Therefore, corresponding to the open neighbourhood $\mathcal{U}$ of $\tilde{\Phi}(y) = \Phi(y)$, there is a chart (W, ρ) at y such that $\tilde{\Phi}(W) \subset \mathcal{U}$ – and then $\Phi(W) \subset \mathcal{V}$.
(iii) As a consequence, the coordinate expressions $\tilde{\Phi}_{\xi\rho}$ and $\Phi_{\eta\rho}$ are both well defined and, since $\xi = pr_1 \circ (\eta \circ \jmath|_{\mathcal{U}})$, it is

$$\tilde{\Phi}_{\xi\rho} = \xi \circ \tilde{\Phi} \circ \rho^{-1} = pr_1 \circ \eta \circ \jmath \circ \tilde{\Phi} \circ \rho^{-1} = pr_1 \circ \eta \circ \Phi \circ \rho^{-1} = pr_1 \circ \Phi_{\eta\rho}$$

This means that $\tilde{\Phi}_{\xi\rho}$ is C^∞ owing to the assumed C^∞-differentiability of $\Phi_{\eta\rho}$. So, for each $y \in N$, there exists a coordinate expression $\tilde{\Phi}_{\xi\rho}$ of $\tilde{\Phi}$ which is C^∞, i.e., $\tilde{\Phi}$ is smooth. ■

Owing to Propositions 3.1.4 and 3.1.9, we know that, on any subset of a manifold, there exists at most one embedded submanifold structure.

3.2 Submersions

Let M, B be smooth manifolds of dimensions m, b respectively, with $m \geq b$ and let

$$\pi : M \longrightarrow B$$

be a smooth mapping. If the tangent map of π at a point $x \in M$ is surjective, π is said to be a *submersion* at x. In such a case, from rank $\pi(x) = \dim \mathrm{Im}\, T_x\pi = \dim T_{\pi(x)}B = b$ it follows that the rank of π is maximal at x. The Rank Theorem 2.2.6 now reads as follows

3.2.1 **Proposition.** *If π is a submersion at x, we can find admissible charts $(\mathcal{U}, \xi)$ and $(\mathcal{V}, \eta)$ at x and $\pi(x)$, respectively, such that*
(o) $\pi(\mathcal{U}) \subset \mathcal{V}$
(i) $\eta \circ \pi|_{\mathcal{U}} = pr_1 \circ \xi$ (where $pr_1 : \mathbb{R}^b \times \mathbb{R}^{m-b} \to \mathbb{R}^b$ is the projection onto the first factor).
Conversely, the existence of charts as above implies that π is a submersion in some open neighbourhood of x. ■

3.2.2 **Existence of a local section.** *Let π be a submersion at x_0. For any open neighbourhood $\mathcal{U}_0$ of x_0, there exists a smooth mapping*

$$\sigma : \mathcal{V}_0 \longrightarrow \mathcal{U}_0$$

defined on an open neighbourhood $\mathcal{V}_0$ of $y_0 := \pi(x_0)$ and taking values in $\mathcal{U}_0$, such that $\sigma(y_0) = x_0$ and

$$\pi \circ \sigma = id_{\mathcal{V}_0} .$$

[15] See Proposition 3.1.1.
[16] See footnote [9].

Proof. Let $(\mathcal{U}, \xi)$ and $(\mathcal{V}, \eta)$ be adapted charts at x_0 and $\pi(x_0)$, respectively. Notice that, up to an intersection with $\mathcal{U}_0$, we can choose ξ with domain $\mathcal{U} \subset \mathcal{U}_0$. Now consider the 'local section' of the natural projection pr_1

$$\rho : \eta(\mathcal{V}) \subset \mathbb{R}^b \longrightarrow \mathbb{R}^b \times \mathbb{R}^{m-b}$$

defined by putting, for each $y \in \mathcal{V}$,

$$\rho\big(\eta(y)\big) = \big(\eta(y), pr_2 \circ \xi(x_0)\big)$$

(where $pr_2 : \mathbb{R}^b \times \mathbb{R}^{m-b} \to \mathbb{R}^{m-b}$ is the projection onto the second factor). As ρ is continuous and

$$\rho\big(\eta(y_0)\big) = \big(\eta \circ \pi(x_0), pr_2 \circ \xi(x_0)\big) = \big(pr_1 \circ \xi(x_0), pr_2 \circ \xi(x_0)\big) = \xi(x_0),$$

corresponding to the open neighbourhood $\xi(\mathcal{U})$ of $\xi(x_0)$, there is an open neighbourhood $\mathcal{V}_0 \subset \mathcal{V}$ of y_0, such that

$$\rho\big(\eta(\mathcal{V}_0)\big) \subset \xi(\mathcal{U}).$$

As a consequence, we can define a smooth mapping $\sigma : \mathcal{V}_0 \to \mathcal{U}_0$ through the composition

$$\sigma := \xi^{-1} \circ \rho \circ \eta\big|_{\mathcal{V}_0}$$

We shall show that σ is the required local section of π. First

$$\sigma(y_0) = \xi^{-1}\big(\rho(\eta(y_0))\big) = \xi^{-1}(\xi(x_0)) = x_0$$

Then, for any $y \in \mathcal{V}_0$,

$$x := \sigma(y) = \xi^{-1}\big(\rho(\eta(y))\big) = \xi^{-1}\big(\eta(y), pr_2 \circ \xi(x_0)\big)$$

on the other hand

$$x = \xi^{-1}\big(\xi(x)\big) = \xi^{-1}\big(pr_1 \circ \xi(x), pr_2 \circ \xi(x)\big) = \xi^{-1}\big(\eta \circ \pi(x), pr_2 \circ \xi(x)\big)$$

hence

$$\begin{aligned}\big(\eta \circ \pi(x), pr_2 \circ \xi(x)\big) = \big(\eta(y), pr_2 \circ \xi(x_0)\big) &\iff \eta \circ \pi(x) = \eta(y) \\ &\iff \pi(x) = y\end{aligned}$$

that is

$$x = \sigma(y) \in \pi^{-1}(y)\,.$$

■

If π is a submersion (i.e., a submersion at every point), the above lemma yields the following

3.2.3 **Theorem.** *A submersion is an open mapping.*

Proof. Let $\pi : M \to B$ be a submersion and $\mathcal{U}_0$ an open subset of M. We will prove that $\pi(\mathcal{U}_0)$ is an open subset of B, by showing that each point $y_0 \in \pi(\mathcal{U}_0)$ (say $y_0 = \pi(x_0)$, with $x_0 \in \mathcal{U}_0$) admits an open neighbourhood $\mathcal{V}_0 \subset \pi(\mathcal{U}_0)$. To this end, it is enough to consider an open neighbourhood $\mathcal{V}_0$ of y_0 where a local section $\sigma : \mathcal{V}_0 \to \mathcal{U}_0$ is defined. For each $y \in \mathcal{V}_0$, we have $y = \pi(\sigma(y))$ with $\sigma(y) \in \mathcal{U}_0$ and then $y \in \pi(\mathcal{U}_0)$. ■

As a consequence, any submersion turns into a surjective one as is proved in the following

3.2.4 **Corollary.** *A submersion induces a submersion onto its own image.*

Proof. Let $\pi : M \to B$ be a submersion and $\tilde{\pi} : M \to \tilde{B}$ the induced mapping onto the image $\tilde{B} := \pi(M)$. Owing to the above theorem, $\tilde{B}$ is an open submanifold of B and then $\tilde{\pi}$ is a smooth mapping [17]. Moreover, since $\pi = \jmath \circ \tilde{\pi}$, the chain rule gives $T_x\pi = T_{\tilde{\pi}(x)}\jmath \circ T_x\tilde{\pi}$ (at each $x \in M$). As a consequence, $T_x\jmath$ being an isomorphism at each $x \in M$, rank $\tilde{\pi}$ = rank $\pi = b = \dim \tilde{B}$. Therefore $\tilde{\pi}$ is a submersion. ∎

Let us now consider two smooth manifolds M, B of dimensions m, b respectively, with $m > b$. Let us consider, for any point $y_0 \in B$, its inverse image $\pi^{-1}(y_0)$ through a smooth mapping $\pi : M \to B$. In the non-trivial case $\pi^{-1}(y_0) \neq \emptyset$, we have the following classical

3.2.5 **Implicit Function Theorem.** *If π is a submersion in $\pi^{-1}(y_0)$, then $\pi^{-1}(y_0)$ is an embedded submanifold of dimension $m - b$.*

Proof. For each $x_0 \in \pi^{-1}(y_0)$, there exist adapted charts (satisfying properties (o), (i) stated in Proposition 3.2.1) $(\mathcal{U}, \xi)$ and $(\mathcal{V}, \eta)$ at x_0 and y_0, respectively. It is

$$\mathcal{U} \cap \pi^{-1}(y_0) = \{x \in \mathcal{U} \,:\, \pi(x) = y_0\}$$

but, if $x \in \mathcal{U}$, we have

$$\begin{aligned} \pi(x) = y_0 &\iff \eta\big(\pi(x)\big) = \eta(y_0) \\ &\iff pr_1\big(\xi(x)\big) = \eta(y_0) =: c \\ &\iff \xi(x) \in \{c\}\times\mathbb{R}^{m-b}. \end{aligned}$$

So the intersection

$$\begin{aligned} \mathcal{U} \cap \pi^{-1}(y_0) &= \{x \in \mathcal{U} \,:\, \xi(x) \in \{c\}\times\mathbb{R}^{m-b}\} \\ &= \xi^{-1}\left(\xi(\mathcal{U}) \cap \left(\{c\}\times\mathbb{R}^{m-b}\right)\right) \\ &= S^{m-b}_{\xi,c}(\mathcal{U}) \end{aligned}$$

is an $(m - b)$-dimensional slice. Now, on the subset of $\mathcal{U}$

$$\tilde{\mathcal{U}} := \mathcal{U} \cap \pi^{-1}(y_0),$$

define the injective map

$$\tilde{\xi} := \alpha \circ \xi|_{\tilde{\mathcal{U}}} : \tilde{\mathcal{U}} \to \mathbb{R}^{m-b},$$

where α is the natural diffeomorphism of $\{c\}\times\mathbb{R}^{m-b}$ onto $\mathbb{R}^{m-b}$. The mapping $\tilde{\xi}$ is an $(m-b)$–chart on $\pi^{-1}(y_0)$, *subordinate* to adapted chart ξ on M. Notice that $\tilde{\xi}$, composition of homeomorphisms, is a homeomorphism. The collection of all such charts on $\pi^{-1}(y_0)$ is an $(m - b)$-dimensional atlas $\mathcal{A}$, whose C^∞-differentiability can be checked as follows. If $\tilde{\xi}' = \alpha \circ \xi|_{\tilde{\mathcal{U}}'}$ and $\tilde{\xi} = \alpha \circ \xi'|_{\tilde{\mathcal{U}}'}$ are charts of $\mathcal{A}$ with non-disjoint domains. either transition function, say

$$\xi' \circ \xi^{-1}\big|_{\xi(\mathcal{U}\cap\mathcal{U}')} = \alpha \circ (\eta' \circ \eta^{-1}) \circ \alpha^{-1}\big|_{\xi(\mathcal{U}\cap\mathcal{U}')}$$

(composition of smooth mappings), is smooth. It follows that, $\pi^{-1}(y_0)$, equipped with the differential structure determined by $\mathcal{A}$, is an embedded submanifold of M, since the manifold topology coincides with subspace topology and at each $x \in \pi^{-1}(y_0)$, one can find a distinguished chart $(\mathcal{U}, \xi)$ on M and the subordinate chart $(\tilde{\mathcal{U}}, \tilde{\xi})$ on $\pi^{-1}(y_0)$, such that the inclusion map $\jmath : \pi^{-1}(y_0) \hookrightarrow M$ satisfies the properties

$$\jmath(\tilde{\mathcal{U}}) \subset \mathcal{U} \quad , \quad \xi \circ \jmath|_{\tilde{\mathcal{U}}} = (\tilde{\xi}, c)\ ,\ c = \text{const}.$$

which make it an immersion [18]. ∎

[17] See 3.1.7 and Proposition 3.1.9.

[18] See Proposition 3.1.1.

3.2.6 **Proposition.** *In the hypothesis of the Implicit Function Theorem, at each point* $x \in \pi^{-1}(y_0)$*, the tangent space* $T_x\big(\pi^{-1}(y_0)\big)$ *is canonically isomorphic to the vector subspace* $\ker T_x\pi \subset T_xM$.

Proof. It is enough to consider (if $\jmath : \pi^{-1}(y_0) \hookrightarrow M$ is the embedding of $\pi^{-1}(y_0)$ into M) the tangent map $T_x\jmath : T_x\pi^{-1}(y_0) \to T_xM$. On the one hand, $T_x\jmath$ is an injective, linear map and then $T_x\big(\pi^{-1}(y_0)\big)$ is isomorphic to Im $T_x\jmath$. On the other hand, from $\pi \circ \jmath = y_0 =$ const., it follows that

$$T_x\pi \circ T_x\jmath = 0$$

that is

$$\mathrm{Im}\ T_x\jmath \subset \ker T_x\pi .$$

Moreover

$$\dim \mathrm{Im}\ T_x\jmath = \dim T_x\big(\pi^{-1}(y_0)\big) = m - b$$
$$\dim \ker T_x\pi = \dim T_xM - \mathrm{rank}\ \pi(x) = m - b$$

and then

$$\mathrm{Im}\ T_x\jmath = \ker T_x\pi .$$

Hence the statement. ∎

3.2.7 **Note.** From the Implicit Function Theorem, we draw that the collection of the *fibres* of a submersion $\pi : M \to B$, i.e., $\big\{\pi^{-1}(y)\,,\, y \in B\big\}$, is a partition of M into $(m-b)$-dimensional, embedded submanifolds. The above partition admits a refiniment given by the collection of the *leaves* of π, i.e., $\big\{L \subset M : L$ is a connected component of a fibre$\big\}$. This is still a partition of M into connected, $(m-b)$-dimensional, embedded submanifolds [(19)].
In general, a k-dimensional *foliation* F of M (with $k < m$) is a partition of M into smoothness preserving, connected, k-dimensional submanifold of M, called leaves of F. If F is a k-dimensional foliation of B, its *lift* by a submersion $\pi : M \to B$, i.e.,

$$\pi^*F := \big\{L \subset M : L \text{ is a connected component of the inverse image } \pi^{-1}(L') \text{ of a leaf } L' \in F\big\}$$

is an $(m-b)+k$-dimensional foliation of M [(20)].

3.3 Local diffeomorphisms

Let M, N be smooth manifolds of dimensions $m = n$ and

$$h : M \longrightarrow N$$

a smooth mapping. If the rank of h is maximal at $x \in M$

$$\mathrm{rank}\ h(x) = m = n,$$

then h is said to be a *local diffeomorphism* at x. In such a case h is both an immersion and a submersion at x, i.e., T_xh is bijective. The Rank Theorem 2.2.6 now reads as follows

(19) Note that any leaf is an open subset of a fibre and then a connected, $(m-b)$-dimensional, embedded submanifold of M (see local connectedness in Sec.1.1, then 3.1.7 and Remark 3.1.2(ii)).

(20) As to the submanifold structure of a leaf and the lift of a foliation, see H.B.Lawson 'Foliations' Bull. Am. Math. Soc., vol.80, n.3 (1974), p.370 and 373. As to the smoothness preserving character of a leaf, see F.Brickell and R.S.Clark, op.cit., p.203.

3.3.1 **Proposition.** *If h is a local diffeomorphism at x, we can find admissible charts $(\mathcal{U}, \xi)$ and $(\mathcal{V}, \eta)$ at x and $h(x)$, respectively, such that*
(o) $h(\mathcal{U}) \subset \mathcal{V}$
(i) $\eta \circ h|_{\mathcal{U}} = \xi$.
On the other hand, the existence of charts as above implies that h is a local diffeomorphism in some open neighbourhood of x. ∎

We can justify the name given to a local diffeomorphism (at a point), by reformulating the above proposition so as to have the following classical

3.3.2 **Inverse Function Theorem.** *A smooth mapping $h : M \to N$ is a local diffeomorphism at x if, and only if, there exixts an open neighbourhood $\mathcal{U}$ of x in M such that $h(\mathcal{U})$ is an open subset of N and h maps $\mathcal{U}$ diffeomorphically onto $h(\mathcal{U})$.*

Proof. Let h be a local diffeomorphism at x. In this case, if $(\mathcal{U}, \xi)$ and $(\mathcal{V}, \eta)$ denote the adapted charts satisfying the above properties (o) and (i), the image $h(\mathcal{U}) = \eta^{-1}\big(\xi(\mathcal{U})\big)$ is an open subset of $\mathcal{V}$ to which, owing to (o), η can be restricted so that we can assume $\mathcal{V} = h(\mathcal{U})$. From (i) we then draw that the induced mapping $\tilde{h}|_{\mathcal{U}} : \mathcal{U} \to h(\mathcal{U})$, is $\eta^{-1} \circ \xi$, which is a diffeomorphism. Conversely, let h be a smooth mapping satisying the requirement stated in the theorem. In this case, the induced diffeomorphism $\tilde{h}|_{\mathcal{U}} : \mathcal{U} \to h(\mathcal{U})$ makes diagram

$$\begin{array}{ccc} \mathcal{U} & \xhookrightarrow{\ \imath\ } & M \\ {\scriptstyle \tilde{h}|_{\mathcal{U}}}\big\downarrow & & \big\downarrow{\scriptstyle h} \\ h(\mathcal{U}) & \xhookrightarrow[\ \jmath\ ]{} & N \end{array}$$

commutative, which implies that $T_x h = T_{h(x)}\jmath \circ T_x \tilde{h}|_{\mathcal{U}} \circ (T_x \imath)^{-1}$ is an isomorphism [21]. ∎

If a discrete topological subspace of a manifold is called embedded submanifold of *zero dimension*, then the above theorem entails the extension of the Implicit Function Theorem to the case $m = n$, as follows

3.3.3 **Corollary.** *If h is a local diffeomorphism in $h^{-1}(y_0)$, then $h^{-1}(y_0)$ is an embedded submanifold of zero dimension.*

Proof. Owing to the Inverse Function Theorem, for each $x \in h^{-1}(y_0)$ there exists an open neighbourhood $\mathcal{U}$ of x such that $h|_{\mathcal{U}}$ is injective. This implies that $\mathcal{U} \cap h^{-1}(y_0) = \{x\}$. Therefore each singleton is an open subset in the subspace topology of $h^{-1}(y_0)$, which is then discrete. ∎

We will now focus on local diffeomorphisms (i.e., mappings which are local diffeomorphisms at every point).

3.3.4 **Corollary.** *A local diffeomorphism is a diffeomorphism if, and only if, it is bijective.*

Proof. A diffeomorphism is, obviously, a bijective local diffeomorphism Conversely, if $h : M \to N$ is a bijective local diffeomorphism, then, owing to the Inverse Function Theorem, it is a smooth mapping with an inverse mapping $h^{-1} : N \to M$ which admits smooth restrictions $\{h^{-1}|_{h(\mathcal{U})}\}$ to suitably many open subsets $\{h(\mathcal{U})\}$ to cover N [22]. ∎

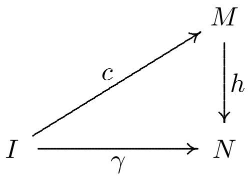
---

[21] See 3.1.7(iii).
[22] See Sec.1.2.

3.3.5 Remark. Any injective local diffeomorphism $h : M \to N$ determines an open submanifold $h(M) \subset N$ diffeomorphic to M (for the induced mapping $\tilde{h} : M \to h(M)$ is a bijective local diffeomorphism). ■

Let $h : M \to N$ be a surjective smooth mapping such that, for each $y \in N$, there exixts a connected open neighbourhood $\mathcal{V}$ of y that is evenly covered by h – i.e., each connected component of $h^{-1}(\mathcal{V})$ is mapped diffeomorphically onto $\mathcal{V}$ by h [23]. We will call h a *covering mapping.* From the inverse function theorem, one immediately draws that a covering mapping is a particular type of surjective local diffeomorphism. It proves to be noticeable in dynamics, owing to the following

3.3.6 Lift theorem. *Let $h : M \to N$ be a covering mapping. If $\gamma : I \to N$ is a smooth curve, then, for any $x_0 \in h^{-1}\big(\gamma(t_0)\big)$, there exists a unique smooth curve $c : I \to M$ that is a lift of γ by h, i.e., $h \circ c = \gamma$,*



with x_0 as initial point at t_0, i.e., $c(t_0) = x_0$.

Proof. Decompose the open interval I into countably many subintervals $\{I_i\}$ such that
(i) $t_0 \in I_0$;
(ii) only any two consecutive subintervals I_i, I_{i+1} have non-empty intersection;
(iii) for any index i, sub-orbit $\gamma(I_i)$ is contained in an evenly covered, connected, open subset $\mathcal{V}_i \subset N$.
Consider the smooth lift of $\gamma|_{I_0}$

$$c_0 := \left(\tilde{h}|_{\mathcal{U}_0}\right)^{-1} \circ \gamma|_{I_0}$$

with initial point x_0. If $t_1 \in I_0 \cap I_1$, repeat the above uniquely determined construction replacing x_0 by $x_1 = c_0(t_1) \in h^{-1}\big(\gamma(t_1)\big)$, so as to have a smooth lift c_1 of $\gamma|_{I_1}$ with initial point x_1. Then, continue by induction on integers i's. Now, any two consecutive lifts c_i, c_{i+1} agree on $I_i \cap I_{i+1}$, in fact, notice that $c_i(I_i \cap I_{i+1})$ is a connected subset of $h^{-1}(\mathcal{V}_{i+1})$ containing x_{i+1} and then contained in $\mathcal{U}_{i+1}$. As a consequence, we can evaluate

$$\left(\tilde{h}|_{\mathcal{U}_{i+1}}\right) \circ c_i|_{I_i \cap I_{i+1}} = h \circ c_i|_{I_i \cap I_{i+1}} = \gamma|_{I_i \cap I_{i+1}} = \left(\tilde{h}|_{\mathcal{U}_{i+1}}\right) \circ c_{i+1}|_{I_i \cap I_{i+1}}$$

whence

$$c_i|_{I_i \cap I_{i+1}} = c_{i+1}|_{I_i \cap I_{i+1}}\,.$$

Therefore $c : I \to M$, defined by $c|_{I_i} = c_i$, is the unique lift we were searching for. ■

---

[23] Recall, e.g., the classical mapping

$$\theta \in \mathbb{R} \mapsto \left(cos\frac{2\pi}{\delta}\,\theta, sin\frac{2\pi}{\delta}\,\theta\right) \in S_1 \qquad (\delta > 0)$$

of $\mathbb{R}$ onto the unit circle S_1 of Euclidean plane $\mathbb{R}^2$.

4 Vector bundles

4.1 Fibre bundles

Let $\pi : M \to B$ be a submersion. As any submersion naturally turns into a surjective one [24], we directly assume π to be surjective. We will also assume $\dim M > \dim B$, in order to have – through the fibres of π – a partition of M into embedded submanifolds [25], whose collection bijectively corresponds to B. In this case, M (*fibred manifold*) can be set-theoretically viewed as the union of disjoint subsets (*fibres*) – each one endowed with a differential structure – whose collection B (*base* or *quotient manifold*) carries a differential structure as well. From the differential structures of the fibres and the base, one obtains the differential structure of M, containing charts which give to any point its coordinates in the fibre where it lies, plus the coordinates of the fibre itself. With this image in mind, π is said to be a (smooth) *fibre bundle*.
Let $\pi : M \to B$ and $\rho : N \to C$ be fibre bundles. A *bundle morphism* from π to ρ is a pair (f, g) of smooth mappings which make the following diagram commutative

$$\begin{array}{ccc} M & \xrightarrow{\quad f \quad} & N \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\rho} \\ B & \xrightarrow[\quad g \quad]{} & C \end{array}$$

Clearly, commutativity property, $\rho \circ f = g \circ \pi$, is completlely equivalent to the *fibre correspondence law*, for any $y \in B$,

$$f\big(\pi^{-1}(y)\big) \subset \rho^{-1}\big(g(y)\big).$$

4.1.1 Proposition. *A bundle morphism induces a smooth mapping between any two fibres which correspond to each other.*

Proof. Just notice that, for any $y \in B$, the restriction $f_y := f|_{\pi^{-1}(y)}$ is a smooth mapping, with values in $\rho^{-1}\big(g(y)\big)$ (embedded submanifold of N). Consequently, owing to Proposition 3.1.9., we have that the induced mapping $\tilde{f}_y : \pi^{-1}(y) \to \rho^{-1}\big(g(y)\big)$ is smooth. ∎

A *bundle isomorphism* from π to ρ is a bundle morphism (f, g) set up by diffeomorphisms. From the above proposition, one immediately draws that

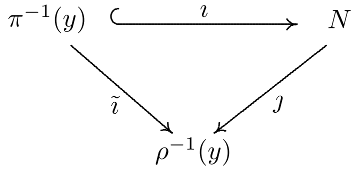
4.1.2 Corollary. *A bundle isomorphism induces a diffeomorphism between any two fibres which correspond to each other.* ∎

Now let $\pi : M \to B$, $\rho : N \to C$ be fibre bundles such that
(i) M is a submanifold of N
(ii) B is a submanifold of C
(iii) the pair of immersions is a bundle morphism from π to ρ.
In such a case, π is said to be a *subbundle* of ρ over B. The name is due to the following

[24] See Corollary 3.2.4.
[25] See Implicit Functions Thorem 3.2.5.

4.1.3 **Proposition.** *If $\pi : M \to B$ is a subbundle of $\rho : N \to C$, then, for each $y \in B$, $\pi^{-1}(y)$ is a submanifold of $\rho^{-1}(y)$.*

Proof. First we notice that $\pi^{-1}(y)$ – as well as $\rho^{-1}(y)$ – is a submanifold of N and $\pi^{-1}(y) \subset \rho^{-1}(y)$. Then, from commutative diagram



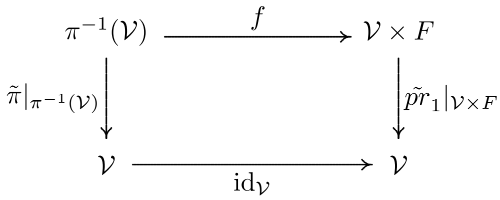
(where $\imath$, $\jmath$ are immersions and $\tilde{\imath}$ is smooth) we draw that

$$\text{rank}\ \imath = \text{rank}\ (\jmath \circ \tilde{\imath}) = \text{rank}\ \tilde{\imath},$$

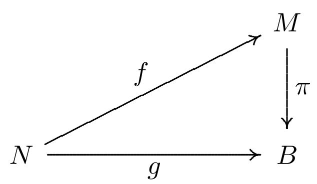
that is, $\tilde{\imath}$ is an immersion. ∎

4.1.4 **Remark.** It is easy to check – through simple set-theoretical considerations on subspace topologies – that, if M is embedded into N, so is $\pi^{-1}(y)$ into $\rho^{-1}(y)$. ∎

Let $\pi : M \to B$ be a fibre bundle, $\mathcal{V}$ an open subset of B and $\pi^{-1}(\mathcal{V})$ the corresponding open subset of M. The induced mapping $\tilde{\pi}|_{\pi^{-1}(\mathcal{V})}$ is a subbundle of π over $\mathcal{V}$, whose fibres are the same as π's over $\mathcal{V}$. Let us now consider a *product bundle* $pr_1 : B \times F \to B$, whose fibres are all canonically diffeomorphic to a given *type fibre* F. For any open subset $\mathcal{V} \subset B$, the above procedure yields a product subbundle of pr_1 over $\mathcal{V}$, $\tilde{pr}_1|_{\mathcal{V}\times F} : \mathcal{V} \times F \to \mathcal{V}$. If there exixts a bundle isomorphism $(f, \text{id}_{\mathcal{V}})$,



then all the fibres of π over $\mathcal{V}$ are diffeomorphic – through f – to the type fibre F. In such a case, f is called a local *trivialization* of π. Most of the fibre bundles we will be dealing with, are *locally trivial* (i.e., each of them admits a local trivialization around each point of its base).
We shall now introduce a particular type of bundle morphism. Let N be a smooth manifold, $\pi : M \to B$ a fibre bundle and (f, g) a bundle morphism from id_N to π. The diagram

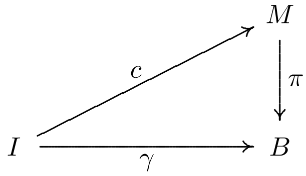


is then commutative, $\pi \circ f = g$, and then, for any $x \in N$,

$$f(x) \in \pi^{-1}\big(g(x)\big)$$

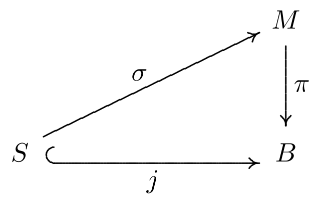
(i.e. $f(x)$ belongs to the fibre of π over $g(x)$). In such a case, f is said to be a *section* of π along g.

Remarkable examples of sections are the following.
(i) Let $I \subset \mathbb{R}$ be an open interval and $\gamma : I \to B$ a smooth curve in B. A section of π along γ is then a smooth curve $c : I \to M$, which projects onto γ by π, i.e.,



$$\pi\big(c(t)\big) = \gamma(t) \qquad \forall t \in I\,.$$

In this case, c is called a *lift* of γ by π.
(ii) Let $j : S \hookrightarrow B$ be the immersion of a submanifold S into B. A section of π along j is then a smooth mapping $\sigma : S \to M$, such that



$$\sigma(x) \in \pi^{-1}(x) \qquad \forall x \in S$$

In this case, σ is called a section of π over $S \subset B$. If S is an open submanifold of B, then σ is a local section of π (26).

Lastly we introduce the following, structurally enriched type of fibre bundle. Let $\pi : M \to B$ be a fibre bundle such that, for each $y \in B$,
(i) $\pi^{-1}(y)$ is a real vector space;
(ii) the embedded submanifold structure of $\pi^{-1}(y)$ coincides with the differential structure determined by the vector space structure.
We will call (M, π, B) a *vector bundle*. A *vector bundle morphism* (between two vector bundles) is a bundle morphism that induces a linear mapping between any two fibres which correspond to each other. A *vector bundle isomorphism* is a vector bundle morphism set up by diffeomorphisms (it induces a linear isomorphism between any two fibres which correspond to each other). A *vector* (resp., *affine*) *subbundle* of a vector bundle π, is a subbundle of π whose fibres are vector (resp., affine) subspaces of the corresponding fibres of π. A vector bundle is locally trivial if, around each point of its base, it admits a local trivialization (onto a product vector bundle) defined by a vector bundle isomorphism.

4.2 Tangent bundle

Let M be an m-dimensional smooth manifold. Recall that, owing to Proposition 2.1.3 and Theorem 2.1.4, for any $x \in M$, the tangent space T_xM, is an m-dimensional vector space and, if $\xi = (x^i)_{i=1,\dots,m}$ is an admissible chart at x, then the coordinate vectors $\big(\frac{\partial}{\partial x^i}\big|_x\big)$ form a basis of T_xM. The corresponding isomorphism $\xi'_x : T_xM \to \mathbb{R}^m$ is given, for each $v \in T_xM$, by

$$\xi'_x(v) = \big(v(x^i)\big)_{i=1,\dots,m}\,.$$

(26) See 3.2.2.

Now let TM be the union of all the tangent spaces of M and

$$\tau_M : TM \to M$$

its natural projection onto M, defined by

$$\tau_M^{-1}(x) := T_x M,$$

for any $x \in M$. Any admissible chart $(\mathcal{U}, \xi)$ on M, determines a $2m$-dimensional *natural chart* $(\mathcal{U}^1, \xi^1)$ on TM defined on

$$\mathcal{U}^1 := \tau_M^{-1}(\mathcal{U})$$

and given by

$$\xi^1 : \mathcal{U}^1 \longrightarrow \xi(\mathcal{U}) \times \mathbb{R}^m \;:\; v = v^i \left.\frac{\partial}{\partial x^i}\right|_x \mapsto \xi^1(v) = \big(\xi(x), \xi'_x(v)\big).$$

Natural charts set up a $2m$-dimensional atlas on TM, whose C^∞-differentiability is shown in the following

4.2.1 **Lemma.** *The atlas of natural charts on TM is C^∞.*

Proof. Let $\xi = (x^1, \ldots, x^m) : \mathcal{U} \to \mathbb{R}^m$ and $\eta = (x^{1'}, \ldots, x^{m'}) : \mathcal{V} \to \mathbb{R}^m$ be admissible charts on M, with $\mathcal{U} \cap \mathcal{V} \neq \emptyset$. If $(\mathcal{U}^1, \xi^1)$ and $(\mathcal{V}^1, \eta^1)$ are the corresponding natural charts on TM, each transition function, say

$$\eta^1 \circ (\xi^1)^{-1} : \xi(\mathcal{U} \cap \mathcal{V}) \times \mathbb{R}^m \longrightarrow \eta(\mathcal{U} \cap \mathcal{V}) \times \mathbb{R}^m \;:\; (\bar{x}, \bar{v}) = (x^i, v^i) \mapsto (x^{j'}, v^{j'})$$

has smooth projections, given (for any $j' = 1, \ldots, m$) by

$$\begin{aligned} x^{j'} &= pr^{j'} \circ \eta^1 \circ \big(\xi^1\big)^{-1}(\bar{x}, \bar{v}) = pr^{j'} \circ \eta \circ \tau_M \left(v^i \left.\frac{\partial}{\partial x^i}\right|_{\xi^{-1}(\bar{x})} \right) \\ &= x^{j'}\big(\xi^{-1}(\bar{x})\big) \end{aligned}$$

and

$$\begin{aligned} v^{j'} &= pr^{m+j'} \circ \eta^1 \circ \big(\xi^1\big)^{-1}(\bar{x}, \bar{v}) = pr^{j'} \circ \eta' \left(v^i \left.\frac{\partial}{\partial x^i}\right|_{\xi^{-1}(\bar{x})} \right) \\ &= \left(v^i \left.\frac{\partial}{\partial x^i}\right|_{\xi^{-1}(\bar{x})} \right) (x^{j'}) \,, \end{aligned}$$

i.e.,

$$v^{j'} = v^i \left.\frac{\partial x^{j'}}{\partial x^i}\right|_{\xi^{-1}(\bar{x})}$$

(contravariant transformation law of the components of a vector). ■

The atlas of natural charts then determines a natural differential structure on TM.

4.2.2 **Proposition.** *If M be a Hausdorff and second-countable, differential manifold, then TM is Hausdorff and second-countable too.*

Proof. First notice that
(i) any two points of TM can be separated by two disjoint coordinate domains;
(ii) TM admit a countable atlas of natural charts.
The thesis then easily follows. ■

We will call τ_M, or TM endowed with its natural differential structure, the *tangent bundle* of M.

4.2.3 Theorem. *$\tau_M : TM \to M$ is a locally trivial vector bundle, with type fibre $\mathbb{R}^m$.*

Proof. For each $x \in M$, let $(\mathcal{U}, \xi)$ be an admissible chart on M at x and $(\mathcal{U}^1, \xi^1)$ the corresponding natural chart on TM.
(i) On the one hand, from the very definition of ξ^1, it follows that τ_M is a submersion and then a fibre bundle.
(ii) On the other hand, the embedded submanifold structure of fibre $\tau_M^{-1}(x)$, is the one containing the global chart ξ'_x subordinate to ξ^1. Also the differential structure determined by the vector structure of $\tau_M^{-1}(x)$ is the one containing the global chart ξ'_x defined by the coordinate vectors. So τ_M is a vector bundle.
(iii) Lastly, check that $(\tilde{\tau}_M|_{\mathcal{U}^1}, \xi^1) : \mathcal{U}^1 \to \mathcal{U} \times \mathbb{R}^m$ is a local trivialization of τ_M at x. ∎

4.2.4 Example. If $M = \mathbb{R}^m$ and $\xi = \mathrm{id}_{\mathbb{R}^m}$, then $\xi^1 = (\tau, \xi')$ is a canonical, global trivialization of $T\mathbb{R}^m$. ∎

Let $\Phi : M \to N$ be a smooth mapping. We will collect all the tangent maps of Φ at the different points of M, into one vector bundle morphism. Define

$$T\Phi : TM \longrightarrow TN$$

by putting, for each $x \in M$ [27],

$$T\Phi|_{T_x M} = T_x\Phi : T_x M \longrightarrow T_{\Phi(x)} N.$$

We call $T\Phi$ the *tangent map* of Φ.

4.2.5 Proposition. *$(T\Phi, \Phi)$ is a vector bundle morphism from τ_M to τ_N.*

Proof. From the very definition of $T\Phi$, it follows that the diagram

$$\begin{array}{ccc} TM & \xrightarrow{\;T\Phi\;} & TN \\ {\scriptstyle\tau_M}\downarrow & & \downarrow{\scriptstyle\tau_N} \\ M & \xrightarrow[\;\Phi\;]{} & N \end{array}$$

is commutative and the induced mapping $T_x\Phi$, between any two fibres which correspond to each other, is linear. So we only have to prove that $T\Phi$ is a smooth mapping. To this end, for any $x \in M$, let $(\mathcal{U}, \xi)$ and $(\mathcal{V}, \eta)$ be admissible charts at x and $\Phi(x)$, with $\Phi(\mathcal{U}) \subset \mathcal{V}$. In the natural charts $(\mathcal{U}^1, \xi^1)$ and $(\mathcal{V}^1, \eta^1)$, the coordinate expression of $T\Phi$

$$\eta^1 \circ T\Phi \circ (\xi^1)^{-1} : \xi(\mathcal{U}) \times \mathbb{R}^m \longrightarrow \eta(\mathcal{V}) \times \mathbb{R}^n \;:\; (\bar{x}, \bar{v}) = (x^i, v^i) \mapsto (y^\alpha, \omega^\alpha)$$

is well defined and has smooth projections, given (for any $\alpha = 1, \ldots, n$) by

$$\begin{array}{ccccc} \mathcal{U}^1 & \xrightarrow{\;T\Phi|_{\mathcal{U}^1}\;} & \mathcal{V}^1 & & \\ {\scriptstyle\xi^1}\downarrow & & \downarrow{\scriptstyle\eta^1} & & \\ \xi(\mathcal{U}) \times \mathbb{R}^m & \xrightarrow[\;(T\Phi)_{\eta^1\xi^1}\;]{} & \eta(\mathcal{V}) \times \mathbb{R}^n & \xrightarrow{\;pr_2\;} & \mathbb{R}^n \\ & & {\scriptstyle pr_1}\downarrow & & \\ & & \eta(\mathcal{V}) & & \end{array}$$

(27) See Sec.2.3.

$$
\begin{aligned}
y^\alpha &= pr^\alpha \circ pr_1 \circ \eta^1 \circ T\Phi \circ \left(\xi^1\right)^{-1}(\bar{x},\bar{v}) = pr^\alpha \circ \eta \circ \tau_N \circ T\Phi \left(v^i \left.\frac{\partial}{\partial x^i}\right|_{\xi^{-1}(\bar{x})} \right) \\
&= y^\alpha \circ \Phi \circ \tau_M \left(v^i \left.\frac{\partial}{\partial x^i}\right|_{\xi^{-1}(\bar{x})} \right) \\
&= y^\alpha \big(\Phi(\xi^{-1}(\bar{x}))\big)\,,
\end{aligned}
$$

and

$$
\begin{aligned}
\omega^\alpha &= pr^\alpha \circ pr_2 \circ \eta^1 \circ T\Phi \circ \left(\xi^1\right)^{-1}(\bar{x},\bar{v}) = pr^\alpha \circ \eta'_{\Phi(x)} \left(T_{\xi^{-1}(\bar{x})}\Phi \cdot v^i \left.\frac{\partial}{\partial x^i}\right|_{\xi^{-1}(\bar{x})} \right) \\
&= \left(T_{\xi^{-1}(\bar{x})}\Phi \cdot v^i \left.\frac{\partial}{\partial x^i}\right|_{\xi^{-1}(\bar{x})} \right)(y^\alpha) \\
&= v^i \left.\frac{\partial(y^\alpha \circ \Phi)}{\partial x^i}\right|_{\xi^{-1}(\bar{x})}
\end{aligned}
$$

The smoothness of all the above functions implies the smoothness of $T\Phi$. ■

Notice that, if Φ is a diffeomorphism, then $T\Phi$ is a diffeomorphism too (for it admits of a smooth inverse mapping) and therefore $(T\Phi, \Phi)$ is a vector bundle isomorphism.

Let X be a differentiable *section* of tangent bundle τ_M over M, i.e., a smooth mapping

$$X : M \longrightarrow TM \;:\; x \mapsto X_x\,,$$

such that

$$\tau_M \circ X = \mathrm{id}_M.$$

The section X is called a *vector field* on M.

4.2.6 Coordinate expression.

Let $(\mathcal{U}, \xi)$ be an admissible chart on M, and $(\mathcal{U}^1, \xi^1)$ the corresponding natural chart on TM. The coordinate expression of the vector field X

$$\xi^1 \circ X \circ \xi^{-1} : \xi(\mathcal{U}) \longrightarrow \xi(\mathcal{U}) \times \mathbb{R}^m$$

has projections given (for any $i = 1, \ldots, m$) by

$$
\begin{aligned}
pr^i \circ \xi^1 \circ X \circ \xi^{-1}(\bar{x}) &= pr^i \circ \xi \circ \tau_M \big(X_{\xi^{-1}(\bar{x})}\big) \\
&= pr^i \circ \xi \circ \xi^{-1}(\bar{x}) \\
&= pr^i(\bar{x})
\end{aligned}
\tag{1}
$$

and

$$
\begin{aligned}
pr^{m+i} \circ \xi^1 \circ X \circ \xi^{-1}(\bar{x}) &= pr^i \circ \xi'_{\xi^{-1}(\bar{x})} \circ X \circ \xi^{-1}(\bar{x}) \\
&= X^i \circ \xi^{-1}(\bar{x})
\end{aligned}
\tag{2}
$$

where

$$X^i : \mathcal{U} \longrightarrow \mathbb{R} \;:\; x \mapsto X^i(x) := X^i_x = X_x(x^i)$$

are the *components* of X in ξ. The first block (1) of projections reduces to $\big(pr^i|_{\xi(\mathcal{U})}\big)$, which are smooth. The second block (2) of projections reduces to $(X^i \circ \xi^{-1})$, which are the coordinate expressions of components (X^i). Therefore, the smoothness of X implies the smothness of components

(X^i) – in any admissible chart. Conversely, the smoothness of components (X^i) – in suitably many admissible charts – ensures the smoothness of X. ∎

We will denote the set of all the vector fields on M by $\chi(M)$. For any $X, Y \in \chi(M)$ and $f \in C^\infty(M)$, the sum $X+Y$ and the multiplication fX (pointwise defined) are both in $\chi(M)$ and give $\chi(M)$ the structure of a $C^\infty(M)$-module. Let W be an open submanifold of M. Owing to the canonical isomorphism $T_xW = T_xM$ (for each $x \in W$), a vector field on W can be viewed as a local vector field on M, i.e., a local section of τ_M, and vice versa. For instance, if $\mathcal{U}$ is the domain of an admissible chart $\xi = (x^i)_{i=1,\dots,m}$ on M, on the one hand, by putting

$$\frac{\partial}{\partial x^i} : x \in \mathcal{U} \longrightarrow \left.\frac{\partial}{\partial x^i}\right|_x \in T\mathcal{U}$$

we have $\frac{\partial}{\partial x^i} \in \chi(\mathcal{U})$, and on the other hand, from any $X \in \chi(M)$, we obtain $X|_\mathcal{U} \in \chi(\mathcal{U})$, related to the previous *coordinate vector fields* by the local decomposition law

$$X|_\mathcal{U} = X^i \frac{\partial}{\partial x^i} .$$

The directional derivative of a function $f \in C^\infty(M)$ along each value of a vector field $X \in \chi(M)$, defines the *Lie derivative* of f along X, given by [(28)]

$$Xf : x \in M \to X_x f \in \mathbb{R}.$$

On the domain $\mathcal{U}$ of any admissible chart, we have

$$(Xf)|_\mathcal{U} = X|_\mathcal{U} f|_\mathcal{U} = \left(X^i \frac{\partial}{\partial x^i} \right) f|_\mathcal{U} = X^i \frac{\partial f}{\partial x^i} \in C^\infty(\mathcal{U})$$

and then $Xf \in C^\infty(M)$. A local example of Lie derivative is

$$X|_\mathcal{U}\, x^i = X^i.$$

The *Lie bracket* of $X, Y \in \chi(M)$, is the vector field

$$[X,Y] : x \in M \mapsto [X,Y]_x \in T_xM \subset TM$$

defined, for any $f \in C^\infty(M)$, by

$$[X,Y]_x f := X_x(Yf) - Y_x(Xf).$$

The Lie derivative along $[X,Y]$ is the commutator of Lie derivatives

$$[X,Y]f = X(Yf) - Y(Xf).$$

On the domain $\mathcal{U}$ of any admissible chart, we have

$$\begin{aligned}[X,Y]^i = [X,Y]_{|\mathcal{U}} x^i &= X_{|\mathcal{U}}(Y_{|\mathcal{U}} x^i) - Y_{|\mathcal{U}}(X_{|\mathcal{U}} x^i) = X_{|\mathcal{U}}(Y^i) - Y_{|\mathcal{U}}(X^i) \\ &= X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \in C^\infty(\mathcal{U})\end{aligned}$$

(28) If W ia an open subset of M, then the Lie derivative along $X \in \chi(M)$ can obviously act on $C^\infty(W)$ by putting, for any $g \in C^\infty(W)$, $Xg := (X|_W)g$.

and then

$$[X, Y] \in \chi(M) \, .$$

One can check that Lie bracket

$$[\ ,\ ] : \chi(M) \times \chi(M) \longrightarrow \chi(M)$$

is an $\mathbb{R}$-bilinear and skew-symmetric operation, satisfying (for any $X, Y, Z \in \chi(M)$) the Jacobi identity

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

Then $\chi(M)$, as a real vector space endowed with the Lie bracket, is a real Lie algebra.

Finally, let $\Phi : M \to N$ be a smooth mapping and $X \in \chi(M)$, $Y \in \chi(N)$. The vector fields X and Y are said to be Φ-related to each other, $X \overset{\Phi}{\sim} Y$, if the diagram

$$\begin{array}{ccc} TM & \xrightarrow{\ T\Phi\ } & TN \\ X \uparrow & & \uparrow Y \\ M & \xrightarrow[\ \Phi\ ]{} & N \end{array}$$

is commutative, i.e.,

$$T\Phi \circ X = Y \circ \Phi.$$

4.2.7 **Lemma.** *The vector fields* X *and* Y *are* Φ*-related to each other if, and only if,* $X(g \circ \Phi) = Yg \circ \Phi$*, for all* $g \in C^\infty(N)$*.*

Proof. For all $x \in M$ and $g \in C^\infty(N)$, it is

$$\begin{aligned} X \overset{\Phi}{\sim} Y &\Longleftrightarrow T\Phi \circ X = Y \circ \Phi \Longleftrightarrow T_x\Phi \cdot X_x = Y_{\Phi(x)} \Longleftrightarrow (T_x\Phi \cdot X_x)g = Y_{\Phi(x)}g \\ &\Longleftrightarrow X_x(g \circ \Phi) = Y_{\Phi(x)}g \Longleftrightarrow \big(X(g \circ \Phi)\big)(x) = \big(Yg \circ \Phi\big)(x) \\ &\Longleftrightarrow X(g \circ \Phi) = Yg \circ \Phi. \end{aligned}$$

■

4.2.8 **Proposition.** *If* $X_1 \overset{\Phi}{\sim} Y_1$ *and* $X_2 \overset{\Phi}{\sim} Y_2$*, then* $[X_1, X_2] \overset{\Phi}{\sim} [Y_1, Y_2]$*.*

Proof. For any $g \in C^\infty(N)$, we have

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

■

4.2.9 **Distributions.** A k-dimensional *distribution* [(29)] on M is a vector subbundle V of TM over M, with k-dimensional fibres $V_x \subset T_xM$. An example is given by the collection of vector subspaces spanned by the values of k $\mathbb{R}$-linearly independent vector fields on M. An *integral manifold* of V is any k-dimensional, connected submanifold $L \hookrightarrow M$ s.t., for each $x \in L$, $T_xj(T_xL) = V_x$. The distribution V is said to be *integrable* if, for each $x \in M$, there exists one, and only one, maximal integral manifold, *leaf*, containing x. In this case, any integral manifold is contained in one leaf and is an open subset of it. The set of all leaves is a k-dimensional foliation of M. A vector field $X \in \chi(M)$ belongs to V if $X_x \in V_x$ for each $x \in M$. The dfistribution V is said to be *involutive* if, for any two nowhere-vanishing vector fields X, Y belonging to V, the commutator $[X, Y]$ belongs to V too. Involutiveness is a necessary and sufficient condition of integrability (Frobenius theorem).

---

(29) See Brickell and R.S.Clark, Differentiable Manifolds, (1970), ch.11.

4.3 Cotangent bundle

Let us consider, at any point $x \in M$, the *cotangent space* $T^*_x M$, dual of $T_x M$, whose elements will be called *covectors* at x. Let $\xi = (x^i)_{i=1,\dots,m}$ be an admissible chart at x. From Remark 2.2.5, it follows that

$$d_x x^i \cdot v = v(x^i) \quad , \quad \forall v \in T_x M.$$

As a consequence, the coordinate covectors $\{d_x x^i\}$ form the dual basis in $T^*_x M$ of basis $\{\frac{\partial}{\partial x^i}|_x\}$ in $T_x M$ which corresponds to the isomorphism

$$\tilde{\xi}_x : T^*_x M \longrightarrow \mathbb{R}^m$$

given, for each $\alpha \in T^*_x M$, by

$$\tilde{\xi}_x(\alpha) := (\alpha_i) = \left(\alpha \cdot \left. \frac{\partial}{\partial x^i} \right|_x \right) .$$

Let T^*M be the (disjoint) union of all the cotangent spaces of M and $\pi_M : T^*M \to M$ its natural projection onto M, defined by

$$\pi_M^{-1}(x) := T^*_x M \qquad \forall x \in M$$

Any m-dimensional admissible chart $(\mathcal{U}, \xi)$ on M, determines a $2m$-dimensional *natural chart* $(\mathcal{U}_1, \xi_1)$ on T^*M defined on

$$\mathcal{U}_1 := \pi_M^{-1}(\mathcal{U}),$$

and given by

$$\xi_1 : \mathcal{U}_1 \longrightarrow \xi(\mathcal{U}) \times \mathbb{R}^m \ : \ \alpha = \alpha_i d_x x^i \mapsto \xi_1(\alpha) = \big(\xi(x), \tilde{\xi}_x(\alpha)\big).$$

Natural charts set up a $2m$-dimensional atlas on T^*M, whose C^∞-differentiability is shown in the following

4.3.1 **Lemma.** *The atlas of natural charts on T^*M is C^∞.*

Proof. Let $\xi = (x^1, \dots, x^m) : \mathcal{U} \to \mathbb{R}^m$ and $\eta = (y^{1'}, \dots, y^{m'}) : \mathcal{V} \to \mathbb{R}^m$, be admissible charts on M, with $\mathcal{U} \cap \mathcal{V} \neq \emptyset$. If $(\mathcal{U}_1, \xi_1)$ and $(\mathcal{V}_1, \eta_1)$ are the corresponding natural charts on T^*M, each transition function, say

$$\eta_1 \circ (\xi_1)^{-1} : \xi(\mathcal{U} \cap \mathcal{V}) \times \mathbb{R}^m \longrightarrow \eta(\mathcal{U} \cap \mathcal{V}) \times \mathbb{R}^m \ : \ (\bar{x}, \bar{\alpha}) = (x^i, \alpha_i) \mapsto (y^{j'}, \alpha_{j'})$$

has smooth projections [(30)], given (for any $j' = 1, \dots, m$) by

$$\begin{aligned} y^{j'} &= pr^{j'} \circ \eta_1 \circ \big(\xi_1\big)^{-1}(\bar{x}, \bar{\alpha}) = pr^{j'} \circ \eta \circ \pi_M \left(\alpha_i d_{\xi^{-1}(\bar{\alpha})} x^i\right) \\ &= y^{j'} \left(\xi^{-1}(\bar{\alpha})\right), \end{aligned}$$

and

$$\begin{aligned} \alpha_{j'} &= pr_{m+j'} \circ \eta_1 \circ \big(\xi_1\big)^{-1}(\bar{x}, \bar{\alpha}) = pr_{j'} \circ \tilde{\eta}_{\xi^{-1}(\bar{x})} \left(\alpha_i d_{\xi^{-1}(\bar{x})} x^i\right) \\ &= \left(\alpha_i d_{\xi^{-1}(\bar{x})} x^i\right) \left(\left. \frac{\partial}{\partial y^{j'}} \right|_{\xi^{-1}(\bar{x})} \right) \end{aligned}$$

(30) As for the smoothness, see the proof of Lemma 4.2.1.

i.e.,

$$\alpha_{j'} = \alpha_i \left.\frac{\partial x^i}{\partial y^{j'}}\right|_{\xi^{-1}(\bar{x})}$$

(covariant transformation law of the components of a covector). ∎

The atlas of natural charts then determines a natural differential structure on T^*M (with a Hausdorff and second-countable topology [(31)]). Endowed with its natural differential structure, T^*M is called the *cotangent bundle* of M.

4.3.2 **Theorem.** *$\pi_M : T^*M \to M$ is a locally trivial vector bundle, with type fibre $\mathbb{R}^m$* [(32)]. ∎

Let θ be a differentiable *section* of the cotangent bundle π_M over M, i.e., a smooth mapping

$$\theta : M \longrightarrow T^*M \ : \ x \mapsto \theta_x \,,$$

such that

$$\pi_M \circ \theta = \mathrm{id}_M .$$

The section θ is called a *covector field* on M.

4.3.3 **Coordinate expression.**
Let $(\mathcal{U}, \xi)$ be an admissible chart on M, and $(\mathcal{U}_1, \xi_1)$ the corresponding natural chart on T^*M. The coordinate expression of the covector field θ,

$$\xi_1 \circ \theta \circ \xi^{-1} : \xi(\mathcal{U}) \longrightarrow \xi(\mathcal{U}) \times \mathbb{R}^m \,,$$

has projections given (for any $i = 1, \ldots, m$) by

$$\begin{aligned} pr^i \circ \xi_1 \circ \theta \circ \xi^{-1}(\bar{x}) &= pr^i \circ \xi \circ \pi_M \big(\theta_{\xi^{-1}(\bar{x})}\big) \\ &= pr^i \circ \xi \circ \xi^{-1}(\bar{x}) \\ &= pr^i(\bar{x}) \end{aligned} \tag{3}$$

and

$$\begin{aligned} pr_{m+i} \circ \xi_1 \circ \theta \circ \xi^{-1}(\bar{x}) &= pr_i \circ \tilde{\xi} \circ \theta \circ \xi^{-1}(\bar{x}) \\ &= \theta_i \circ \xi^{-1}(\bar{x}) \end{aligned} \tag{4}$$

where

$$\theta_i : \mathcal{U} \longrightarrow \mathbb{R} \ : \ x \mapsto \theta_i(x) := (\theta_x)_i = \theta_x \left.\frac{\partial}{\partial x^i}\right|_x$$

are the components of θ in ξ. The first block (3) of projections reduces to $\big(pr^i|_{\xi(\mathcal{U})}\big)$, which are smooth. The second block (4) of projections reduces to $(\theta_i \circ \xi^{-1})$, which are the coordinate expressions of components (θ_i). Therefore, the smoothness of θ implies the smothness of the components (θ_i) – in any admissible chart. Conversely, the smoothness of the components (θ_i) – in suitably many admissible charts – ensures the smoothness of θ. ∎

(31) See Proposition 4.2.2.
(32) See the proof of Theorem 4.2.3.

4.3.4 **Example.** Let $f \in C^\infty(M)$. The differential df [33] is the covector field

$$df : x \in M \longrightarrow d_x f \in T_x^* M \subset T^* M,$$

whose smoothness is ensured by the smoothness of its components in any admissible chart $(\mathcal{U}, \xi)$, given by

$$(df)_i : \mathcal{U} \longrightarrow \mathbb{R} \, : \, x \mapsto d_x f \left(\left. \frac{\partial}{\partial x^i} \right|_x \right) = \left. \frac{\partial f}{\partial x^i} \right|_x$$

that is

$$(df)_i = \frac{\partial f}{\partial x^i} \, . \qquad \blacksquare$$

We will denote the set of all the covector fields on M by $\chi^*(M)$. For any $\theta, \theta' \in \chi^*(M)$ and $f \in C^\infty(M)$, the sum $\theta + \theta'$ and the multiplication $f\theta$ (pointwise defined) are both in $\chi^*(M)$ and give $\chi^*(M)$ the structure of a $C^\infty(M)$-module. Let W be an open submanifold of M. Owing to the canonical isomorphism $T_x^* W = T_x^* M$ (for each $x \in W$), a covector field on W can be viewed as a local covector field on M, i.e., a local section of π_M, and vice versa. For instance, if $\mathcal{U}$ is the domain of an admissible chart $\xi = (x^i)_{i=1,\ldots,m}$ on M, on the one hand, by putting

$$dx^i : x \in \mathcal{U} \longrightarrow d_x x^i \in T^* \mathcal{U}$$

we have $dx^i \in \chi^*(\mathcal{U})$, and on the other hand, from any $\theta \in \chi^*(M)$, we obtain $\theta|_\mathcal{U} \in \chi^*(\mathcal{U})$, related to the previous *coordinate covector fields* by the local decomposition law

$$\theta|_\mathcal{U} = \theta_i dx^i \, .$$

The pointwise interaction between a covector field $\theta \in \chi^*(M)$ and a vector field $X \in \chi(M)$, defines the *pairing*

$$\theta X : x \in M \longrightarrow \theta_x X_x \in \mathbb{R} \, .$$

On the domain $\mathcal{U}$ of an admissible chart ξ, we have

$$\begin{aligned} (\theta X)|_\mathcal{U} &= \theta|_\mathcal{U} X|_\mathcal{U} = (\theta_i dx^i) \left(X^j \frac{\partial}{\partial x^j} \right) = \theta_i X^j \left(dx^i \frac{\partial}{\partial x^j} \right) = \theta_i X^j \delta^i_j \\ &= \theta_i X^i \in C^\infty(\mathcal{U}) \end{aligned}$$

and then

$$\theta X \in C^\infty(M) \, .$$

An example of pairing is the Lie derivative

$$Xf = df X \, .$$

Local examples of pairings are the components of a vector or covector field

$$X^i = dx^i \, X|_\mathcal{U}$$

$$\theta_i = \theta|_\mathcal{U} \, \frac{\partial}{\partial x^i} \, .$$

[33] See Remark 2.2.5(i).

4.4 Tensor bundles

In this section we will embody tangent vectors and covectors in the general context of tensor calculus.

Let $x \in M$ and let r, s be non-negative integers. For $r = s = 0$, put

$$(T_0^0 M)_x = \{x\} \times \mathbb{R}$$

and, in any other case, let $(T_s^r M)_x$ denote the set of all the $\mathbb{R}$-multilinear forms of the following type

$$A : \underbrace{T_x^*M \times \ldots \times T_x^*M}_{r \text{ times}} \times \underbrace{T_xM \times \ldots \times T_xM}_{s \text{ times}} \longrightarrow \mathbb{R}.$$

Notice that, for $r = 0$ and $s = 1$, we have, up to a canonical isomorphism,

$$(T_1^0 M)_x = T_x^* M$$

and , for $r = 1$ and $s = 0$, we have

$$(T_0^1 M)_x = T_x^{**} M = T_x M\,.$$

In any case, $(T_s^r M)_x$ is called the $\binom{r}{s}$*-tensor space* at x, its elements being called (r times contravariant and s times covariant) *tensors* at x. Sum and multiplication by real numbers are natural operations in $(T_s^r M)_x$, which is so given the structure of a vector space. Another useful operation on tensors is the *tensor product*, which acts on a pair $A \in (T_s^r M)_x$, $B \in (T_{s'}^{r'} M)_x$ and yields $A \otimes B \in (T_{s+s'}^{r+r'} M)_x$ given by ordinary multiplication AB if $r = s = 0$ or $r' = s' = 0$, and, in any other case, by

$$A \otimes B\,(\alpha^1, \ldots \alpha^r, \beta^1, \ldots \beta^{r'}, u_1, \ldots u_s, v_1, \ldots v_{s'}) := A\,(\alpha^1, \ldots \alpha^r, u^1, \ldots u^s)\, B\,(\beta^1, \ldots \beta^{r'}, v^1, \ldots v^{s'})$$

In particular, if $\xi = (x^i)_{1,\ldots,m}$ is an admissible chart at x, then the system of m^{r+s} *coordinate tensor products*

$$\left(\frac{\partial}{\partial x^{i_1}}\bigg|_x \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}}\bigg|_x \otimes d_x x^{j_1} \otimes \ldots \otimes d_x x^{j_s} \right) \tag{5}$$

is a basis of $(T_s^r M)_x$, as is shown in the following

4.4.1 **Proposition.** *Any tensor $A \in (T_s^r M)_x$ can be uniquely expressed as a linear combination*

$$A = A_{j_1 \ldots j_s}^{i_1 \ldots i_r} \frac{\partial}{\partial x^{i_1}}\bigg|_x \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}}\bigg|_x \otimes d_x x^{j_1} \otimes \ldots \otimes d_x x^{j_s}$$

with components

$$A_{j_1 \ldots j_s}^{i_1 \ldots i_r} := A \left(d_x x^{i_1}, \ldots, d_x x^{i_r} | \frac{\partial}{\partial x^{j_1}}\bigg|_x, \ldots, \frac{\partial}{\partial x^{j_s}}\bigg|_x \right) \,.$$

Proof. (i) For each

$$(\alpha^1, \ldots \alpha^r, u^1, \ldots u^s) \in \underbrace{T_x^*M \times \ldots \times T_x^*M}_{r \text{ times}} \times \underbrace{T_xM \times \ldots \times T_xM}_{s \text{ times}}$$

we have

$$\begin{aligned}
&A\,(\alpha^1,\dots\alpha^r,u^1,\dots u^s) = \\
&= A\left(\alpha^1_{i_1}\,d_x x^{i_1},\dots,\alpha^r_{i_r}\,d_x x^{i_r},u^{j_1}_1\,\frac{\partial}{\partial x^{j_1}}\bigg|_x,\dots,u^{j_s}_s\,\frac{\partial}{\partial x^{j_s}}\bigg|_x\right) \\
&= A^{i_1\dots i_r}_{j_1\dots j_s}\;\alpha^1_{i_1}\dots\alpha^r_{i_r}\;u^{j_1}_1\dots u^{j_s}_s \\
&= A^{i_1\dots i_r}_{j_1\dots j_s}\;\alpha^1\left(\frac{\partial}{\partial x^{i_1}}\bigg|_x\right)\dots\alpha^r\left(\frac{\partial}{\partial x^{i_r}}\bigg|_x\right)d_x x^{j_1}(u_1)\dots d_x x^{j_s}(u_s) \\
&= \left(A^{i_1\dots i_r}_{j_1\dots j_s}\,\frac{\partial}{\partial x^{i_1}}\bigg|_x\otimes\dots\otimes\frac{\partial}{\partial x^{i_r}}\bigg|_x\otimes d_x x^{j_1}\otimes\dots\otimes d_x x^{j_s}\right)(\alpha^1,\dots\alpha^r,u^1,\dots u^s),
\end{aligned}$$

whence the required decomposition of A.
(ii) The uniqueness of the above decomposition is due to the linear independence of the coordinate tensor products. This can be checked by evaluating a vanishing linear combination

$$A^{i_1\dots i_r}_{j_1\dots j_s}\left(\frac{\partial}{\partial x^{i_1}}\bigg|_x\otimes\dots\otimes\frac{\partial}{\partial x^{i_r}}\bigg|_x\otimes d_x x^{j_1}\otimes\dots\otimes d_x x^{j_s}\right)=0$$

on

$$\left(d_x x^{h_1},\dots,d_x x^{h_r},\frac{\partial}{\partial x^{k_1}}\bigg|_x,\dots,\frac{\partial}{\partial x^{k_s}}\bigg|_x\right).$$

We obtain

$$A^{i_1\dots i_r}_{j_1\dots j_s}\,\delta^{h_1}_{i_1}\dots\delta^{h_r}_{i_r}\,\delta^{j_1}_{k_1}\dots\delta^{j_s}_{k_s}=0$$

that is

$$A^{h_1\dots h_r}_{k_1\dots k_s}=0$$

for all the values of the indexes. ■

Owing to the above proposiition, $\dim\left(T^r_sM\right)_x = m^{r+s}$ and basis (5) of $\left(T^r_sM\right)_x$ corresponds to the isomorphism

$$\bar{\xi}_x:\left(T^r_sM\right)_x\longrightarrow\mathbb{R}^{m^{r+s}}$$

given, for any $A\in\left(T^r_sM\right)_x$, by

$$\bar{\xi}_x(A):=\left(A\left(d_x x^{i_1},\dots,d_x x^{i_r},\frac{\partial}{\partial x^{j_1}}\bigg|_p,\dots,\frac{\partial}{\partial x^{j_s}}\bigg|_x\right)\right)$$

(for $r=s=0$, just put $\bar{\xi}_x:=\mathrm{id}_{\mathbb{R}}$).
Now let T^r_sM be the (disjoint) union of all the $\binom{r}{s}$-tensor spaces of M. Notice that

$$T^0_0M=M\times\mathbb{R}\quad,\quad T^1_0M=TM\quad,\quad T^0_1M=T^*M.$$

Let

$$\tau^r_{s\,M}:T^r_sM\longrightarrow M$$

be the natural projection of T^r_sM onto M, defined by

$$(\tau^r_{s\,M})^{-1}(x):=\left(T^r_sM\right)_x\qquad\forall x\in M.$$

Any admissible chart $(\mathcal{U},\xi)$ on M, determines a $(m+m^{r+s})$-*natural chart* on T^r_sM defined on

$$\mathcal{U}^r_s := (\tau^r_{s\,M})^{-1}(\mathcal{U})$$

and given by

$$\xi^r_s : \mathcal{U}^r_s \longrightarrow \xi(\mathcal{U})\times\mathbb{R}^{m^{r+s}}$$

$$A = A^{i_1\dots i_r}_{j_1\dots j_s}\left(\left.\frac{\partial}{\partial x^{i_1}}\right|_x \otimes\dots\otimes \left.\frac{\partial}{\partial x^{i_r}}\right|_x \otimes d_xx^{j_1}\otimes\dots\otimes d_xx^{j_s}\right) \mapsto \xi^r_s(A) = \big(\xi(x),\bar{\xi}_x(A)\big)\,.$$

Natural charts set up a $(m+m^{r+s})$-dimensional atlas on T^r_sM, whose C^∞-differentiability is shown in the following

4.4.2 Lemma. *The atlas of natural charts on T^r_sM is C^∞.*

Proof. Let $\xi = (x^1,\dots,x^m) : \mathcal{U} \to\mathbb{R}^m$ and $\eta = (y^{1'},\dots,y^{m'}) : \mathcal{V} \to\mathbb{R}^m$ be admissible charts on M, with $\mathcal{U}\cap\mathcal{V}\neq\emptyset$. If $(\mathcal{U}^r_s,\xi^r_s)$ and $(\mathcal{V}^r_s,\eta^r_s)$ are the corresponding natural charts on T^r_sM, each transition function, say

$$\eta^r_s\circ(\xi^r_s)^{-1} : \xi(\mathcal{U}\cap\mathcal{V})\times\mathbb{R}^{m^{r+s}} \longrightarrow \eta(\mathcal{U}\cap\mathcal{V})\times\mathbb{R}^{m^{r+s}} \;:\; (\bar{x},\bar{A}) = \big(x^i, A^{i_1\dots i_r}_{j_1\dots j_s}\big) \mapsto \Big(y^{j'},\; A^{i'_1\dots i'_r}_{j'_1\dots j'_s}\Big)$$

has smooth projections [34], given (for any $j'=1,\dots,m$) by

$$\begin{aligned} y^{j'} &= u^{j'}\circ pr_1\circ\eta^r_s\circ\big(\xi^r_s\big)^{-1}(\bar{x},\bar{A})\\ &= u^{j'}\circ\eta\circ\sigma\left(A^{i_1\dots i_r}_{j_1\dots j_s}\left(\left.\frac{\partial}{\partial x^{i_1}}\right|_x\otimes\dots\otimes\left.\frac{\partial}{\partial x^{i_r}}\right|_x\otimes d_xx^{j_1}\otimes\dots\otimes d_xx^{j_s}\right)\right)\\ &= y^{j'}\big(\xi^{-1}(\bar{x})\big)\,, \end{aligned}$$

and (for any $i'_1\dots i'_r\,,\, j'_1\dots j'_s = 1,\dots,m$)

$$\begin{aligned} A^{i'_1\dots i'_r}_{j'_1\dots j'_s} &= u^{i'_1\dots i'_r}_{j'_1\dots j'_s}\circ pr_2\circ\eta^r_s\circ\big(\xi^r_s\big)^{-1}(\bar{x},\bar{A})\\ &= u^{i'_1\dots i'_r}_{j'_1\dots j'_s}\circ\bar{\eta}_{\xi^{-1}(\bar{x})}\left(A^{i_1\dots i_r}_{j_1\dots j_s}\left.\frac{\partial}{\partial x^{i_1}}\right|_{\xi^{-1}(\bar{x})}\otimes\dots\otimes\left.\frac{\partial}{\partial x^{i_r}}\right|_{\xi^{-1}(\bar{x})}\otimes d_{\xi^{-1}(\bar{x})}x^{j_1}\otimes\dots\otimes d_{\xi^{-1}(\bar{x})}x^{j_s}\right)\\ &= \left(A^{i_1\dots i_r}_{j_1\dots j_s}\left.\frac{\partial}{\partial x^{i_1}}\right|_{\xi^{-1}(\bar{x})}\otimes\dots\otimes\left.\frac{\partial}{\partial x^{i_r}}\right|_{\xi^{-1}(\bar{x})}\otimes d_{\xi^{-1}(\bar{x})}x^{j_1}\otimes\dots\otimes d_{\xi^{-1}(\bar{x})}x^{j_s}\right)\\ &\qquad\left(d_{\xi^{-1}(\bar{x})}y^{j'_1}\otimes\dots\otimes d_{\xi^{-1}(\bar{x})}y^{j'_s}\otimes\left.\frac{\partial}{\partial y^{i'_1}}\right|_{\xi^{-1}(\bar{x})}\otimes\dots\otimes\left.\frac{\partial}{\partial y^{i'_r}}\right|_{\xi^{-1}(\bar{x})}\right)\\ &= A^{i_1\dots i_r}_{j_1\dots j_s}\left.\frac{\partial y^{i'_1}}{\partial x^{i_1}}\right|_{\xi^{-1}(\bar{x})}\cdots\left.\frac{\partial y^{i'_s}}{\partial x^{i_r}}\right|_{\xi^{-1}(\bar{x})}\left.\frac{\partial x^{j_1}}{\partial y^{j'_1}}\right|_{\xi^{-1}(\bar{x})}\cdots\left.\frac{\partial x^{j_s}}{\partial y^{j'_r}}\right|_{\xi^{-1}(\bar{x})} \end{aligned}$$

(r-times contravariant and s-times covariant transformation law of the components of a tensor). ∎

The atlas of natural charts then determines a natural differential structure on T^r_sM (with a Hausdorff and second-countable topology [35]). Endowed with its natural differential structure, $\tau^r_{s\,M}$, or T^r_sM, is called the $\binom{r}{s}$- *tensor bundle* of M.

---

(34) As for the smoothness, see the proof of Lemma 4.2.1.
(35) See Proposition 4.2.2.

4.4.3 **Theorem.** *$\tau^r_{s\,M} : T^r_s M \to M$ is a locally trivial vector bundle, with type fibre $\mathbb{R}^{m^{r+s}}$* [36]. ∎

Let A be a differentiable section of the $\binom{r}{s}$-tensor bundle $\tau^r_{s\,M}$ over M, i.e., a smooth mapping

$$A : M \to T^r_s M$$

such that

$$\tau^r_{s\,M} \circ A = \mathrm{id}_M \,.$$

The section A is called a $\binom{r}{s}$-*tensor field* on M.

4.4.4 **Coordinate expression.**
Let $(\mathcal{U}, \xi)$ be an admissible chart on M, and $(\mathcal{U}^s_r, \xi^s_r)$ the corresponding natural chart on $T^r_s M$. The coordinate expression of a tensor field A,

$$\xi^r_s \circ A \circ \xi^{-1} : \xi(\mathcal{U}) \longrightarrow \xi(\mathcal{U}) \times \mathbb{R}^{m^{r+s}},$$

has projections given (for any $i = 1, \ldots, m$) by

$$\begin{aligned} u^i \circ pr_1 \circ \xi^r_s \circ A \circ \xi^{-1}(\bar{x}) &= u^i \circ \xi \circ \sigma\big(A_{\xi^{-1}(\bar{x})}\big) \\ &= u^i \circ \xi \circ \xi^{-1}(\bar{x}) \\ &= u^i(\bar{x}) \end{aligned} \tag{6}$$

and (for any $i_1, \ldots, i_r, j_1, \ldots, j_s = 1, \ldots, m$) by

$$\begin{aligned} u^{i_1 \ldots i_r}_{j_1 \ldots j_s} \circ pr_2 \circ \xi^r_s \circ A \circ \xi^{-1}(\bar{x}) &= u^{i_1 \ldots i_r}_{j_1 \ldots j_s} \circ \bar{\xi} \circ A \circ \xi^{-1}(\bar{x}) \\ &= A^{i_1 \ldots i_r}_{j_1 \ldots j_s} \circ \xi^{-1}(\bar{x}) \end{aligned} \tag{7}$$

where

$$\begin{aligned} A^{i_1 \ldots i_r}_{j_1 \ldots j_s} : \mathcal{U} \longrightarrow \mathbb{R} \ : \ x \mapsto A^{i_1 \ldots i_r}_{j_1 \ldots j_s}(x) :&= (A_x)^{i_1 \ldots i_r}_{j_1 \ldots j_s} \\ &= A_x\left(d_x x^{i_1}, \ldots, d_x x^{i_r}, \left.\frac{\partial}{\partial x^{j_i}}\right|_x, \ldots, \left.\frac{\partial}{\partial x^{j_s}}\right|_x\right) \end{aligned}$$

are the *components* of A in ξ. The first block (6) of projections reduces to $\big(u^i|_{\xi(\mathcal{U})}\big)$, which are smooth. The second block (7) of projections reduces to $(A^{i_1 \ldots i_r}_{j_1 \ldots j_s} \circ \xi^{-1})$, which are the coordinate expressions of A's components. Therefore, the smoothness of A implies the smothness of its components – in any admissible chart. Conversely, the smoothness of A's components – in suitably many admissible charts – ensures the smoothness of A. ∎

[36] See the proof of Theorem 4.2.3.

4.4.5 **Example.** Let A, B be tensor fields of type $\binom{r}{s}$ and $\binom{r'}{s'}$ respectively. Their tensor product

$$A \otimes B : M \longrightarrow T^{r+r'}_{s+s'}M \ : \ x \in M \mapsto A_x \otimes B_x \in \left(T^{r+r'}_{s+s'}M\right)_x$$

is an $\binom{r+r'}{s+s'}$-tensor field, whose smoothness is ensured by the smoothness of its components, given (in any admissible chart) by

$$(A \otimes B)^{i_1 \ldots i_r i_{r+1} \ldots i_{r+r'}}_{j_1 \ldots j_s j_{s+1} \ldots j_{s+s'}} := A^{i_1 \ldots i_r}_{j_1 \ldots j_s} B^{i_{r+1} \ldots i_{r+r'}}_{j_{s+1} \ldots j_{s+s'}}$$ ■

We will denote the set of all the $\binom{r}{s}$-tensor fields on M by $\chi^r_s(M)$.
Notice the following identifications:

$$\chi^0_0(M) = C^\infty(M) \ , \ \chi^1_0(M) = \chi(M) \ , \ \chi^0_1(M) = \chi^*(M) \ .$$

For any A, $A' \in \chi^r_s(M)$ and $f \in C^\infty(M)$, the sum $A + A'$ and the multiplication fA (pointwise defined) are both in $\chi^r_s(M)$ and give $\chi^r_s(M)$ the structure of a $C^\infty(M)$-module.
Let W be an open submanifold of M. Owing to the canonical isomorphism $(T^r_s W)_x = (T^r_s M)_x$ (for each $x \in W$), a tensor field on W can be viewed as a local tensor field on M, i.e., a local section of ${\tau^r_s}_M$, and vice versa. For instance, if $\mathcal{U}$ is the domain of an admissible chart $\xi = (x^i)_{i=1,\ldots,m}$ on M, on the one hand, we have

$$\frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s} \in \chi^r_s(\mathcal{U})$$

and, on the other hand, from any $A \in \chi^r_s(M)$, we obtain $A|_{\mathcal{U}} \in \chi^r_s(\mathcal{U})$, related to the previous *coordinate tensor fields* by the local decomposition law

$$A|_{\mathcal{U}} = A^{i_1 \ldots i_r}_{j_1 \ldots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s} \ .$$

The pointwise action of a tensor field $A \in \chi^r_s M$ on covector and vector fields defines the *pairing*

$$\tilde{A} : \underbrace{\chi^*(M) \times \ldots \times \chi^*(M)}_{r \text{ times}} \times \underbrace{\chi(M) \times \ldots \times \chi(M)}_{s \text{ times}} \longrightarrow C^\infty(M) \tag{8}$$

given by

$$\tilde{A}(\theta^1, \ldots, \theta^r, X_1, \ldots, X_s)(x) := A_x(\theta^1_x, \ldots, \theta^r_x, X_{1x}, \ldots, X_{sx}) \ .$$

Notice that $\tilde{A}$ is well defined since on the domain of any admissible chart $(\mathcal{U}, \xi)$, we have

$$\tilde{A}(\theta^1, \ldots, \theta^r, X_1, \ldots, X_s)|_{\mathcal{U}} = A^{i_1 \ldots i_r}_{j_1 \ldots j_s} \theta^1_{i_1}, \ldots, \theta^r_{i_r} X^{j_1}_1, \ldots, X^{j_s}_s \in C^\infty(\mathcal{U})$$

and then

$$\tilde{A}(\theta^1, \ldots, \theta^r, X_1, \ldots, X_s) \in C^\infty(M).$$

Moreover $\tilde{A}$ is easily checked to be $C^\infty(M)$-multilinear. Let us denote by $\tilde{\chi}^r_s(M)$ the $C^\infty(M)$-module of all the $C^\infty(M)$-multilinear mappings of type (8).

4.4.6 **Theorem.** *The mapping*

$$\tilde{} : \chi^r_s(M) \longrightarrow \tilde{\chi}^r_s(M) \;:\; A \mapsto \tilde{A}$$

is an isomorphism.

Proof. For the sake of notational simplicity, we shall put $r = 0$ and $s = 1$ and then we shall prove the bijectivity of the mapping

$$\tilde{}: \chi^*(M) \longrightarrow \tilde{\chi}^*(M)$$

given by

$$\tilde{\theta}(X) := \theta X.$$

Let $\tilde{\theta}: \chi(M) \to C^\infty(M)$ be an $C^\infty(M)$-linear mapping.
(o) Let $X \in \chi M$ be such that $X_x = 0$. We shall show that $(\tilde{\theta}X)_x = 0$. To this purpose, let $(\mathcal{U}, \xi)$ be an admissible chart at x, where $X|_{\mathcal{U}} = X^i \frac{\partial}{\partial x^i}$, and $\beta \in C^\infty(M)$ a bump function at x with support in $\mathcal{U}$. As is known [37], we can 'extend' both $X^i \in C^\infty(\mathcal{U})$ and $\frac{\partial}{\partial x^i} \in \chi(\mathcal{U})$ to $\tilde{X}^i \in C^\infty(M)$ and $\frac{\tilde{\partial}}{\partial x^i} \in \chi(M)$ by putting

$$\tilde{X}^i|_{\mathcal{U}} = \beta|_{\mathcal{U}} X^i \quad , \quad \tilde{X}^i|_{M-\mathcal{U}} = 0$$

$$\frac{\tilde{\partial}}{\partial x^i}\Big|_{\mathcal{U}} = \beta|_{\mathcal{U}} \frac{\partial}{\partial x^i} \quad , \quad \frac{\tilde{\partial}}{\partial x^i}\Big|_{M-\mathcal{U}} = 0.$$

From

$$\beta^2 X = \tilde{X}^i \frac{\tilde{\partial}}{\partial x^i}$$

we draw

$$\beta^2(\tilde{\theta}X) = \tilde{\theta}(\beta^2 X) = \tilde{\theta}\left(\tilde{X}^i \frac{\tilde{\partial}}{\partial x^i}\right) = \tilde{X}^i \left(\tilde{\theta} \frac{\tilde{\partial}}{\partial x^i}\right).$$

Evaluating the left and right hand side at x (where $\beta(x) = 1$ and $\tilde{X}^i(x) = X^i(x) = 0$), we have $(\tilde{\theta}X)(x) = 0$. As a consequence, if $X, Y \in \chi(M)$ are such that $X_x = Y_x$, i.e.,

$$(X - Y)(x) = 0,$$

we have

$$0 = \big(\tilde{\theta}(X - Y)\big)(x) = (\tilde{\theta}X - \tilde{\theta}Y)(x) = (\tilde{\theta}X)(x) - (\tilde{\theta}Y)(x)$$

that is $(\tilde{\theta}X)(x) = (\tilde{\theta}Y)(x)$. So the value of $\tilde{\theta}X$ at x only depends on the value of X at x.
(i) Owing to (o), we can define

$$\theta : M \longrightarrow T^*M$$

by putting, for any $x \in M$ and $v \in T_xM$,

$$\theta_x(v) := (\tilde{\theta}X)(x)$$

---

[37] See 1.3.

where $X \in \chi(M)$ is any vector field such that $X_x = v$ [38]. Clearly $\theta_x : T_xM \to \mathbb{R}$ is linear, i.e., $\theta_x \in T_x^*M$. The smoothness of θ is ensured by smoothness ot its components (θ_i) in any admissible chart $(\mathcal{U}, \xi)$, for (if $\mathcal{V} \subset \mathcal{U}$ denotes an open neighbourhood of a $x \in \mathcal{U}$ where $\beta|_{\mathcal{V}} = 1$)

$$\theta|_{\mathcal{V}} \left.\frac{\partial}{\partial x^i}\right|_{\mathcal{V}} = \theta_i|_{\mathcal{V}} = \left.\left(\tilde{\theta}\frac{\tilde{\partial}}{\partial x^i}\right)\right|_{\mathcal{V}} \in C^\infty(\mathcal{V}).$$

So $\theta \in \chi^*(M)$. Since

$$\theta X = \tilde{\theta} X \quad , \quad \forall X \in \chi(M), \tag{9}$$

$\tilde{\theta}$ is the image of θ under $\sim$. This shows that $\sim$ is surjective.
(ii) $\tilde{\theta}$ is the image of just one $\theta \in \chi^*(M)$ under $\sim$. Indeed, owing to (9), there is no choice as to the value θ_x of θ at any $x \in M$, θ_x is to be the linear form acting on each $v \in T_xM$ as follows:

$$\theta_x v = \theta_x X_x = (\theta X)(x) = (\tilde{\theta} X)(x),$$

where, $X \in \chi(M)$ is any vector field such that $X_x = v$. This shows that $\sim$ is injective. ∎

Owing to Theorem 4.4.6, a tensor field $A \in \chi^r_s(M)$ will now be meant as an $C^\infty(M)$-multilinear mapping

$$A : \underbrace{\chi^*(M) \times \ldots \times \chi^*(M)}_{r\,\text{times}} \times \underbrace{\chi(M) \times \ldots \times \chi(M)}_{s\,\text{times}} \longrightarrow C^\infty(M).$$

Restriction of A to any s-tuple $X_1, \ldots, X_s \in \chi(M)$, is a $C^\infty(M)$-multilinear mapping

$$A(\underbrace{\ldots}_{r}; X_1, \ldots, X_s) : \underbrace{\chi^*(M) \times \ldots \times \chi^*(M)}_{r\,\text{times}} \longrightarrow C^\infty(M)$$

and then

$$A(\underbrace{\ldots}_{r}; X_1, \ldots, X_s) \in \chi^r_0(M).$$

The consequent mapping

$$\hat{A} : \underbrace{\chi(M) \times \ldots \times \chi(M)}_{s\,\text{times}} \longrightarrow \chi^r_0(M) \tag{10}$$

defined by

$$\hat{A}(X_1, \ldots, X_s) := A(\underbrace{\ldots}_{r}; X_1, \ldots, X_s) \tag{11}$$

is easily checked to be $C^\infty(M)$-multilinear.
Let us denote by $\hat{\chi}^r_s(M)$ the $C^\infty(M)$-module of all the $C^\infty(M)$-multilinear mappings of type (10).

---

[38] For example, if $(\mathcal{U}, \xi)$ is an admissible chart at x, consider $Y \in \chi(\mathcal{U})$ characterized by constant components $Y^i = v^i$ in ξ; then 'extend' Y to an $X \in \chi(M)$ by means of am bump function β at x with support in $\mathcal{U}$.

4.4.7 **Theorem.** *The mapping*

$$\hat{} : \chi^r_s(M) \longrightarrow \hat{\chi}^r_s(M) \ : \ A \mapsto \hat{A}$$

is an isomorphism.

Proof. Let $\hat{A}$ be an $C^\infty(M)$-multilinear mapping of type (10).
(i) We obtain an $A \in \chi^r_s(M)$, by putting, for each $\theta^1, \dots, \theta^r \in \chi^*(M)$ and $X_1, \dots, X_s \in \chi(M)$,

$$A(\theta^1, \dots, \theta^r, X_1, \dots, X_s) := \hat{A}(X_1, \dots, X_s)(\theta^1, \dots, \theta^r), \tag{12}$$

and $\hat{A}$ is the image of A under $\hat{}$. This shows that $\hat{}$ is surjective.
(ii) $\hat{A}$ is the image of just one $A \in \chi^r_s(M)$ under $\hat{}$. Indeed, owing to (11), there is no choice as to the value of A at any $\theta^1, \dots, \theta^r \in \chi^*(M)$ and $X_1, \dots, X_s \in \chi(M)$; this value is to be (12). This shows that $\hat{}$ is injective. ∎

Owing to Theorem 4.4.7, a tensor field $A \in \chi^r_s(M)$ can also be meant as an $C^\infty(M)$-multilinear mapping

$$A : \underbrace{\chi(M) \times \dots \times \chi(M)}_{s\,\text{times}} \longrightarrow \chi^r_0(M)$$

4.4.8 **Remark.** On the domain of any admissible chart $(\mathcal{U}, \xi)$, from the local decomposition

$$A|_\mathcal{U} = A^{i_1 \dots i_r}_{j_1 \dots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s} \,,$$

we draw

$$\begin{aligned} A|_\mathcal{U}\left(\frac{\partial}{\partial x^{j_1}}, \dots, \frac{\partial}{\partial x^{j_s}}\right) &= \left[A|_\mathcal{U}\left(\frac{\partial}{\partial x^{j_1}}, \dots, \frac{\partial}{\partial x^{j_s}}\right)\right]^{i_1 \dots i_r} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_r}} \\ &= \left[A|_\mathcal{U}\left(dx^{i_1}, \dots, dx^{i_r}, \frac{\partial}{\partial x^{j_1}}, \dots, \frac{\partial}{\partial x^{j_s}}\right)\right] \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_r}} \\ &= A^{i_1 \dots i_r}_{j_1 \dots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_r}} \end{aligned}$$

and then the above local decomposition reads

$$A|_\mathcal{U} = A|_\mathcal{U}\left(\frac{\partial}{\partial x^{j_1}}, \dots, \frac{\partial}{\partial x^{j_s}}\right) \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s}.$$ ∎

5 External forms and derivations

5.1 External forms

Let M be an m-dimensional smooth manifold and s be an integer. If $s < 0$, let $\Lambda_s M$ be the null $C^\infty(M)$-module. If $s = 0$, let

$$\Lambda_0 M := \chi_0^0 M = C^\infty(M)$$

be the $C^\infty(M)$-module of $\binom{0}{0}$-tensor fields, i.e., real smooth functions, on M. If $s = 1$, let

$$\Lambda_1 M := \chi_1^0 M = \chi^* M$$

be the $C^\infty(M)$-module of $\binom{0}{1}$-tensor fields. If $s > 1$, let

$$\Lambda_s M \subset \chi_s^0 M$$

be the $C^\infty(M)$-module of skew-symmetric, $\binom{0}{s}$-tensor fields on M. Any such field, as a $C^\infty(M)$-multilinear mapping

$$\omega : \underbrace{\chi M \times \ldots \times \chi M}_{s\,\text{times}} \longrightarrow C^\infty(M),$$

satisfies, for all $X_1, \ldots, X_s \in \chi M$ and $i, j = 1, \ldots, s$, the *skew-symmetry condition*

$$\omega(X_1, \ldots, X_i, \ldots, X_j, \ldots, X_s) = -\omega(X_1, \ldots, X_j, \ldots, X_i, \ldots, X_s).$$

Notice that, for any skew-symmetric $\omega \in \Lambda_s M$, the restriction $\omega|_{\mathcal{U}}$ to the domain of any admissible chart $(\mathcal{U}, \xi)$, is skew-symmetric too and then, if $s > m$, $\omega|_{\mathcal{U}}\left(\frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_s}}\right) = 0$. From the decomposition law of $\omega|_{\mathcal{U}}$ it follows that $\omega|_{\mathcal{U}} = 0$ and then, from the arbitrariness of ξ, we have $\omega = 0$. So, also for $s > m$, $\Lambda_s M = \{0\}$.
For any integer s, the elements of $\Lambda_s M$ are called *exterior s-differential forms*. In the set ΛM of all the exterior forms on M, called *Grassman* or *exterior algebra* of M, the *exterior product*

$$\wedge : \Lambda M \times \Lambda M \longrightarrow \Lambda M$$

is an (associative, $C^\infty(M)$-bilinear) operation that maps each $\tau \in \Lambda_r M$ and $\omega \in \Lambda_s M$ onto an image $\tau \wedge \omega \in \Lambda_{s+r} M$ given by

$$\tau \wedge \omega = 0$$

if either r or s is negative, and, in any other case, by

$$\begin{aligned} \tau \wedge \omega\big(X_1, \ldots, X_{s+r}\big) = &= \frac{1}{s!r!} \sum_\sigma (\operatorname{sign} \sigma) \tau \otimes \omega\big(X_{\sigma(1)}, \ldots, X_{\sigma(r+s)}\big) \\ &= \frac{1}{s!r!} \sum_\sigma (\operatorname{sign} \sigma) \tau\big(X_{\sigma(1)}, \ldots, X_{\sigma(r)}\big) \omega\big(X_{\sigma(r+1)}, \ldots, X_{\sigma(r+s)}\big) \end{aligned}$$

(the sum being extended to all the permutations σ of $\{1, \ldots, s+r\}$). Hence

$$\tau \wedge \omega = (-1)^{rs} \omega \wedge \tau.$$

5.1.1 Examples. Check that
(i) if $f \in \Lambda_0 M$ and $\omega \in \Lambda_s M$ with $s \geq 0$, then

$$f \wedge \omega = f\omega\,;$$

(ii) if $\tau, \omega \in \Lambda_1 M$, then

$$\tau \wedge \omega = \tau \otimes \omega - \omega \otimes \tau\,.$$

5.1.2 Coordinate expression.
In any admissible chart $(\mathcal{U}, \xi)$, the components of $\tau \wedge \omega \in \Lambda_{s+r} M$ are

$$\begin{aligned}(\tau \wedge \omega)_{j_1 \ldots j_{r+s}} &= (\tau \wedge \omega)|_{\mathcal{U}} \left(\frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_{r+s}}} \right) \\ &= \frac{1}{r!s!} \sum_{\sigma} (\operatorname{sign} \sigma) \tau_{j_{\sigma(1)} \cdots j_{\sigma(r)}} \omega_{j_{\sigma(r+1)} \cdots j_{\sigma(r+s)}}\end{aligned}$$

In particular, for $r = s = 1$, $(\tau \wedge \omega)_{ij} = \tau_i \omega_j - \tau_j \omega_i$.

5.1.3 Remark. The local decomposition law of any $\omega \in \Lambda_s M$, with $s > 1$, in an admissible chart $(\mathcal{U}, \xi)$, reads

$$\omega|_{\mathcal{U}} = \sum_{j_1 < \ldots < j_s} \omega_{j_1 \ldots j_s} dx^{j_1} \wedge \ldots \wedge dx^{j_s}\,.$$

Proof. For the sake of simplicity, we shall put $s = 2$. The local decomposition law of $\omega \in \Lambda_2 M$ in $(\mathcal{U}, \xi)$, is

$$\omega|_{\mathcal{U}} = \omega_{ij} dx^i \otimes dx^j\,,$$

with

$$\omega_{ij} = \omega|_{\mathcal{U}} \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right).$$

Owing to the skew-symmetry of $\omega|_{\mathcal{U}}$, we have

$$\omega_{ij} = \begin{cases} -\omega_{ji} & i \neq j \\ 0 & i = j \end{cases}$$

and then

$$\begin{aligned}\omega|_{\mathcal{U}} &= \sum_{i<j} (\omega_{ij} dx^i \otimes dx^j + \omega_{ji} dx^j \otimes dx^i) = \sum_{i<j} \omega_{ij} (dx^i \otimes dx^j - dx^j \otimes dx^i) \\ &= \sum_{i<j} \omega_{ij} dx^i \wedge dx^j\,.\end{aligned}$$

■

Let $\Phi : M \to N$ be a smooth mapping. The *pull-back* by Φ is the additive operator

$$\Phi^* \colon \Lambda N \longrightarrow \Lambda M$$

which maps $\omega \in \Lambda_s N$ onto $\Phi^* \omega \in \Lambda_s M$ given by

$$\Phi^* \omega = \begin{cases} 0 & s < 0 \\ \omega \circ \Phi & \mathrm{s} = 0 \end{cases}$$

and for $s > 0$,

$$(\Phi^* \omega)_x = \omega_{\Phi(x)} \circ (\underbrace{T_x \Phi \times \ldots \times T_x \Phi}_{s \text{ times}})$$

for all $x \in M$.

5.1.4 **Coordinate expression.**
Let $\xi = (x^j)_{j=1,\dots,m} : \mathcal{U} \to \mathbb{R}^m$ and $\eta = (y^\alpha)_{\alpha=1,\dots,n} : \mathcal{V} \to \mathbb{R}^n$, with $\Phi(\mathcal{U}) \subset \mathcal{V}$, be admissible charts on M and N, respectively. Recall [39] that, at any $x \in \mathcal{U}$, the relation $w = T_x\Phi \cdot v$ is expressed in ξ and η by $w = w^\alpha \left.\frac{\partial}{\partial y^\alpha}\right|_{\Phi(x)}$ with $w^\alpha = \left.\frac{\partial \Phi^\alpha}{\partial x^i}\right|_x v^i$ (where $\Phi^\alpha := y^\alpha \circ \Phi|_{\mathcal{U}}$). As a consequence,

$$T_x\Phi \cdot \left.\frac{\partial}{\partial x^j}\right|_x = \left.\frac{\partial \Phi^\alpha}{\partial x^j}\right|_x \left.\frac{\partial}{\partial y^\alpha}\right|_{\Phi(x)} .$$

Now let $\omega \in \Lambda_s N$, with $s > 0$. At any $x \in \mathcal{U}$, we have

$$\begin{aligned}
(\Phi^*\omega)_{j_1\dots j_s}(x) = (\Phi^*\omega)_{x\, j_1\dots j_s} &= (\Phi^*\omega)_x \left(\left.\frac{\partial}{\partial x^{j_1}}\right|_x, \dots, \left.\frac{\partial}{\partial x^{j_s}}\right|_x \right) \\
&= \omega_{\Phi(x)} \left(T_x\Phi \cdot \left.\frac{\partial}{\partial x^{j_1}}\right|_x, \dots, T_x\Phi \cdot \left.\frac{\partial}{\partial x^{j_s}}\right|_x \right) \\
&= \omega_{\Phi(x)} \left(\left.\frac{\partial \Phi^{\alpha_1}}{\partial x^{j_1}}\right|_x \left.\frac{\partial}{\partial y^{\alpha_1}}\right|_{\Phi(x)}, \dots, \left.\frac{\partial \Phi^{\alpha_s}}{\partial x^{j_s}}\right|_x \left.\frac{\partial}{\partial y^{\alpha_s}}\right|_{\Phi(x)} \right) \\
&= \left.\frac{\partial \Phi^{\alpha_1}}{\partial x^{j_1}}\right|_x \cdots \left.\frac{\partial \Phi^{\alpha_s}}{\partial x^{j_s}}\right|_x \omega_{\Phi(x)} \left(\left.\frac{\partial}{\partial y^{\alpha_1}}\right|_{\Phi(x)}, \dots, \left.\frac{\partial}{\partial y^{\alpha_s}}\right|_{\Phi(x)} \right) \\
&= \left.\frac{\partial \Phi^{\alpha_1}}{\partial x^{j_1}}\right|_x \cdots \left.\frac{\partial \Phi^{\alpha_s}}{\partial x^{j_s}}\right|_x \omega_{\Phi(x)\, \alpha_1\dots\alpha_s} \\
&= \left.\frac{\partial \Phi^{\alpha_1}}{\partial x^{j_1}}\right|_x \cdots \left.\frac{\partial \Phi^{\alpha_s}}{\partial x^{j_s}}\right|_x \omega_{\alpha_1\dots\alpha_s}(\Phi(x)) .
\end{aligned}$$

So $\Phi^*\omega$ has components

$$(\Phi^*\omega)_{j_1\dots j_s} = \frac{\partial \Phi^{\alpha_1}}{\partial x^{j_1}} \cdots \frac{\partial \Phi^{\alpha_s}}{\partial x^{j_s}} (\omega_{\alpha_1\dots\alpha_s} \circ \Phi|_{\mathcal{U}})$$

(whose smoothness ensures the smoothness of $\Phi^*\omega$). ■

5.1.5 **Proposition.** *The pull back operator Φ^* preserves the exterior product, i.e.,*

$$\Phi^*(\tau \wedge \omega) = \Phi^*\tau \wedge \Phi^*\omega ,$$

for any τ, $\omega \in \Lambda N$.

Proof. For the sake of notational simplicity, we shall put τ, $\omega \in \Lambda_1 N$.
On the one hand, we have

$$\big(\Phi^*(\tau \wedge \omega)\big)_{i,j} = \frac{\partial \Phi^\alpha}{\partial x^i}\frac{\partial \Phi^\beta}{\partial x^j}(\tau \wedge \omega)_{\alpha\beta} \circ \Phi|_{\mathcal{U}} = \frac{\partial \Phi^\alpha}{\partial x^i}\frac{\partial \Phi^\beta}{\partial x^j}(\tau_\alpha\omega_\beta - \tau_\beta\omega_\alpha) \circ \Phi|_{\mathcal{U}} .$$

On the other hand, from

$$\begin{aligned}
(\Phi^*\tau)_i &= \frac{\partial \Phi^\alpha}{\partial x^i} (\tau_\alpha \circ \Phi|_{\mathcal{U}}) \\
(\Phi^*\omega)_j &= \frac{\partial \Phi^\beta}{\partial x^j} (\omega_\beta \circ \Phi|_{\mathcal{U}}),
\end{aligned}$$

[39] See 2.2.4.

it follows that

$$\begin{aligned}(\Phi^*\tau\wedge\Phi^*\omega)_{ij} &= (\Phi^*\tau)_i(\Phi^*\omega)_j-(\Phi^*\tau)_j(\Phi^*\omega)_i\\ &= \frac{\partial\Phi^\alpha}{\partial x^i}\frac{\partial\Phi^\beta}{\partial x^j}(\tau_\alpha\omega_\beta\circ\Phi|_{\mathcal U})-\frac{\partial\Phi^\beta}{\partial x^j}\frac{\partial\Phi^\alpha}{\partial x^i}(\tau_\beta\omega_\alpha\circ\Phi|_{\mathcal U})\\ &= \frac{\partial\Phi^\alpha}{\partial x^i}\frac{\partial\Phi^\beta}{\partial x^j}(\tau_\alpha\omega_\beta-\tau_\beta\omega_\alpha)\circ\Phi|_{\mathcal U}\,.\end{aligned}$$

Hence the statement. ■

5.1.6 **Remark.** The action of Φ^* – as above defined between ΛN and ΛM – naturally extends to the action between $\Lambda\mathcal V$ and $\Lambda\mathcal U$ where $\mathcal V$ is any open submanifold of N and $\mathcal U:=\Phi^{-1}(\mathcal V)$ its open inverse image in M. From this extension of Φ^*, it follows that, for any $\omega\in\Lambda N$,

$$\Phi^*(\omega|_{\mathcal V})=(\Phi^*\omega)|_{\mathcal U}\,.$$

5.2 Derivations

Let k be an integer. A *derivation* of *degree* k on the exterior algebra ΛM, after Frolicher and Nijenhuis [40] is an operator

$$D:\Lambda M\to\Lambda M$$

that maps each $\omega\in\Lambda_s M$ onto an image $D\omega\in\Lambda_{s+k}M$, satisfying (for each τ, $\omega\in\Lambda M$ and a, $b\in\mathbb R$)
(i) $\mathbb R$-linearity

$$D(a\tau+b\omega)=aD\tau+bD\omega$$

(in every $\Lambda_s M$);
(ii) Leibniz rule

$$D(\tau\wedge\omega)=D\tau\wedge\omega+(-1)^{rk}\tau\wedge D\omega$$

(where r is the degree of τ).
The local character of a derivation is pointed out in the following

5.2.1 **Proposition.** *If σ, $\tau\in\Lambda_r M$ coincide on an open subset $\mathcal U\subset M$,*

$$\sigma|_{\mathcal U}=\tau|_{\mathcal U}\,,$$

then

$$(D\sigma)|_{\mathcal U}=(D\tau)|_{\mathcal U}\,.$$

Proof. If we put $\omega:=\sigma-\tau$, we have $\omega|_{\mathcal U}=0$. Then, if $x\in\mathcal U$ and β is a bump function at x with support in $\mathcal U$, we also have $\omega=(1-\beta)\omega$ both on $\mathcal U$ (where ω vanishes) and on $M-\mathcal U$ (where β vanishes). Now, on the one hand, Leibniz rule implies

$$D\omega=D(1-\beta)\wedge\omega+(1-\beta)D\omega$$

[40] See A.Frolicher and A.Nijenhuis, Theory of vector-valued differential forms, Ind.Math., 18 (1956), p.338-385.

and then

$$(D\omega)_x = 0$$

since $\omega_x = 0$ and $\beta(x) = 1$. On the other hand, $\mathbb{R}$-linearity implies

$$D\omega = D\sigma - D\tau$$

and then

$$(D\sigma)_x = (D\tau)_x \, .$$

Hence, owing to the arbitrariness of $x \in \mathcal{U}$, the statement follows. ∎

A consequence of the 'locality' of a derivation D on ΛM, is the possibility of extending the action of D to the exterior algebra ΛW of any open submanifold $W \subset M$. To this end, let $\alpha \in \Lambda W$. If $\sigma, \tau \in \Lambda M$ are equal to α around a point $x \in W$, i.e.,

$$\sigma|_{\mathcal{V}_1} = \alpha|_{\mathcal{V}_1} \quad , \quad \tau|_{\mathcal{V}_2} = \alpha|_{\mathcal{V}_2}$$

on the open neighbourhoods $\mathcal{V}_1$, $\mathcal{V}_2 \subset W$ of x, then

$$\sigma|_{\mathcal{V}_1 \cap \mathcal{V}_2} = \alpha|_{\mathcal{V}_1 \cap \mathcal{V}_2} = \tau|_{\mathcal{V}_1 \cap \mathcal{V}_2}$$

and, owing to the above proposition,

$$(D\sigma)_x = (D\tau)_x \, .$$

Therefore the action of D can be extended to α by putting, at each $x \in W$,

$$(D\alpha)_x := (D\sigma)_x$$

with any $\sigma \in \Lambda M$ equal to α around x (the smoothness of $D\alpha$ is ensured by the smoothness of its components, which, in a suitably small coordinate neighbourhood of x, are the components of $D\sigma$). Clearly this action of D on ΛW still satisfies $\mathbb{R}$-linearity and Leibniz rule, i.e., D is a derivation on ΛW too. The above extension clearly implies that, for any open submanifold $W \subset M$ and any $\omega \in \Lambda M$,

$$D(\omega|_W) = (D\omega)|_W \, .$$

5.2.2 **Lemma.** *Let D and D' be derivations on ΛM. If for each $f \in C^\infty(M)$,*

$$Df = D'f$$
$$Ddf = D'df \, ,$$

then

$$D = D' \, .$$

Proof.
(i) Let us first show that, if W is an open submanifold of M, then, for each $\varphi \in C^\infty(W)$,

$$D\varphi = D'\varphi \quad , \quad Dd\varphi = D'd\varphi \, .$$

To this purpose, let $x \in W$. If $f \in C^\infty(M)$ is equal to φ around x, i.e., $f|_\mathcal{V} = \varphi|_\mathcal{V}$ on an open neighbourhood $\mathcal{V} \subset W$ of x, then df is equal to $d\varphi$ on $\mathcal{V}$, and consequently

$$(D\varphi)_x := (Df)_x = (D'f)_x =: (D'\varphi)_x \,,$$

$$(Dd\varphi)_x := (Ddf)_x = (D'df)_x =: (D'd\varphi)_x \,.$$

(ii) Let us now prove the main statement, i.e.,

$$D\omega = D'\omega \quad , \quad \forall \omega \in \Lambda M.$$

For the sake of simplicity, as usual, we shall only check the result for $\omega \in \Lambda_1 M$. On the domain of any admissible chart $(\mathcal{U}, \xi)$ on M, we have

$$(D\omega)|_\mathcal{U} = D(\omega|_\mathcal{U}) = D(\omega_i dx^i) = D\omega_i \wedge dx^i + \omega_i D dx^i$$

Then, owing to (i),

$$(D\omega)|_\mathcal{U} = D'\omega_i \wedge dx^i + \omega_i D' dx^i = D'(\omega_i dx^i) = D'(\omega|_\mathcal{U}) = (D'\omega)|_\mathcal{U}$$

whence the statement. ∎

Let D_1 and D_2 be derivations on ΛM of degree k_1 and k_2, respectively. The operator

$$[D_1, D_2] : \Lambda M \longrightarrow \Lambda M$$

defined by

$$[D_1, D_2] := D_1 D_2 - (-1)^{k_1 k_2} D_2 D_1$$

is called the *commutator* of D_1 and D_2.

5.2.3 **Remark.** Notice that

$$-(-1)^{k_1 k_2}[D_1, D_2] = -(-1)^{k_1 k_2} D_1 D_2 + D_2 D_1 = [D_2, D_1].$$ ∎

5.2.4 **Proposition.** *The operator $[D_1, D_2]$ is a derivation of degree $k_1 + k_2$.*

Proof. From the definition of commutator, one can directly draw that $[D_1, D_2]$ takes, $\mathbb{R}$-linearly, any $\Lambda_s M$ into $\Lambda_{s+(k_1+k_2)} M$. Then we only have to prove Leibniz rule

$$\underbrace{[D_1, D_2](\sigma \wedge \tau)}_{(1)} = \underbrace{\big([D_1, D_2]\sigma\big) \wedge \tau + (-1)^{(k_1+k_2)s} \sigma \wedge \big([D_1, D_2]\tau\big)}_{(2)}$$

for any $\sigma \in \Lambda_s M, \tau \in \Lambda M$. As to right hand side (2), we have

$$\begin{aligned}
(2) &= \big(D_1 D_2 \sigma - (-1)^{k_1 k_2} D_2 D_1 \sigma\big) \wedge \tau \\
&\quad + (-1)^{(k_1+k_2)s} \sigma \wedge \big(D_1 D_2 \tau - (-1)^{k_1 k_2} D_2 D_1 \tau\big) \\
&= \underbrace{\big(D_1 D_2 \sigma\big) \wedge \tau}_{1} \underbrace{- (-1)^{k_1 k_2} \big(D_2 D_1 \sigma\big) \wedge \tau}_{2} \\
&\quad + \underbrace{(-1)^{(k_1+k_2)s} \sigma \wedge \big(D_1 D_2 \tau\big)}_{3} \underbrace{- (-1)^{(k_1+k_2)s} (-1)^{k_1 k_2} \sigma \wedge \big(D_2 D_1 \tau\big)}_{4} .
\end{aligned}$$

As to left hand side (1), we have

$$(1) = D_1 D_2(\sigma \wedge \tau) - (-1)^{k_1 k_2} D_2 D_1(\sigma \wedge \tau)$$

where the first term of the sum is

$$\begin{aligned} D_1\big(D_2\sigma \wedge \tau + (-1)^{sk_2}\sigma \wedge D_2\tau\big) &= D_1\big(D_2\sigma \wedge \tau\big) + (-1)^{sk_2} D_1\big(\sigma \wedge D_2\tau\big) \\ &= \underbrace{\big(D_1 D_2\sigma\big) \wedge \tau}_{1} + \underbrace{(-1)^{(s+k_2)k_1} D_2\sigma \wedge D_1\tau}_{a} \\ &\quad + \underbrace{(-1)^{sk_2} D_1\sigma \wedge D_2\tau}_{b} + \underbrace{(-1)^{sk_2+sk_1}\sigma \wedge \big(D_1 D_2\tau\big)}_{3} \end{aligned}$$

and the second term (which can be obtained from the first one via permutation of $\{1,2\}$ and multiplication by $-(-1)^{k_1 k_2}$) is

$$\begin{aligned} &\underbrace{-(-1)^{k_1 k_2}\big(D_2 D_1\sigma\big) \wedge \tau}_{2}\underbrace{-(-1)^{k_1 k_2+(s+k_1)k_2} D_1\sigma \wedge D_2\tau}_{-b} \\ &\underbrace{-(-1)^{k_1 k_2+sk_1} D_2\sigma \wedge D_1\tau}_{-a}\underbrace{-(-1)^{k_1 k_2+sk_1+sk_2}\sigma \wedge \big(D_2 D_1\tau\big)}_{4}. \end{aligned}$$

So (with the above numeration) we have

$$(1) = (1 + a + b + 3) + (2 - b - a + 4) = 1 + 2 + 3 + 4 = (2)$$

that is the statement. ∎

5.3 Cartan calculus

A classical derivation on ΛM, which extends the ordinary differentiation on $C^\infty(M)$ (and is at the foundation of the so called Cartan calculus), is given in the following

5.3.1 **Theorem.** *There exists a unique derivation d of degree 1 on ΛM, called exterior derivative, such that, for any $f \in C^\infty(M)$,*
(i) df is the differential of f,
(ii) $d(df) = 0$.
(iii) The action of d vanishes on $\Lambda_s M$ for $s < 0$, is given by (i) on $\Lambda_0 M$, and, on each $\omega \in \Lambda_s M$ with $s > 0$, is given by

$$\begin{aligned} d\omega(X_0, X_1, \dots, X_s) := &\sum_{i=0}^{s} (-1)^i X_i \omega(X_0, \dots, \hat{X}_i, \dots, X_s) \\ &+ \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_s) \end{aligned} \tag{13}$$

(where $X_0, X_1, \dots, X_s \in \chi M$ and symbol $\hat{}$ denotes omission of the term where it is placed).

Proof. Existence. One should prove that operator d, defined by (13) is a derivation which satisfies condition (ii). We shall only check (ii). If $\omega \in \Lambda_1 M$, then, for any $X_0 = X, X_1 = Y \in \chi M$, we have

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

If $\omega = df$ (with $f \in C^\infty(M)$),

$$ddf(X, Y) = X(Yf) - Y(Xf) - [X, Y]f = 0.$$

Hence the statement.
The uniqueness follows from the above Lemma 5.2.2. ∎

5.3.2 **Remark.** If you want to be led to find out action (13) of exterior derivative d (i.e., of a derivation of degree 1 satisfying (i), (ii)), you can follow these directions. First, check that d satisfies (i), (ii) on any open submanifold $\mathcal{U} \subset M$ too. Then, let $\mathcal{U}$ be the domain of an admissible chart. On $\mathcal{U}$, for any $\omega \in \Lambda M$, e.g. $\omega \in \Lambda_1 M$, we have

$$\begin{aligned}(d\omega)|_{\mathcal{U}} &= d(\omega|_{\mathcal{U}}) = d(\omega_j dx^j) = d\omega_j \wedge dx^j = \frac{\partial \omega_j}{\partial x^i} dx^i \wedge dx^j \\ &= \frac{\partial \omega_j}{\partial x^i}(dx^i \otimes dx^j - dx^j \otimes dx^i)\end{aligned}$$

or, equivalently,

$$(d\omega)|_{\mathcal{U}} = \left(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j}\right) dx^i \otimes dx^j \,.$$

As a consequence, for any $X, Y \in \chi M$, it is

$$\begin{aligned}\big(d\omega(X,Y)\big)|_{\mathcal{U}} &= d(\omega|_{\mathcal{U}})(X|_{\mathcal{U}}, Y|_{\mathcal{U}}) = \left(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j}\right) X^i Y^j = X^i \frac{\partial \omega_j}{\partial x^i} Y^j - Y^j \frac{\partial \omega_i}{\partial x^j} X^i \\ &= X^i \frac{\partial}{\partial x^i}(\omega_j Y^j) - Y^j \frac{\partial}{\partial x^j}(\omega_i X^i) - X^i \frac{\partial Y^j}{\partial x^i}\omega_j + Y^j \frac{\partial X^i}{\partial x^j}\omega_i \\ &= X^i \frac{\partial}{\partial x^i}(\omega_j Y^j) - Y^j \frac{\partial}{\partial x^j}(\omega_i X^i) - \omega_j \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i}\right) \\ &= \big(X(\omega Y) - Y(\omega X) - \omega[X,Y]\big)|_{\mathcal{U}}.\end{aligned}$$

On M then we have

$$d\omega(X_0, X_1) = X_0(\omega X_1) - X_1(\omega X_0) - \omega[X_0, X_1]$$

(with $X_0 = X$ and $X_1 = Y$), that is

$$d\omega(X_0, X_1) = \sum_{i=0}^{1}(-1)^i X_i \omega(\hat{X}_i, X_{i+1}) + \sum_{i<j}(-1)^{i+j}\omega([X_i, X_j], \hat{X}_i, \hat{X}_j).$$

■

5.3.3 **Property.** *The exterior derivative is nilpotent, i.e.,*

$$d^2 = 0.$$

Proof. The proof is based on the following remark: Leibniz rule and condition (ii) imply that

$$d(d\varphi_1 \wedge \ldots \wedge d\varphi_s) = 0$$

for any $\varphi_1, \ldots, \varphi_s \in C^\infty(\mathcal{U})$ and any open submanifold $\mathcal{U} \subset M$. With this remark, we evaluate d on $\omega \in \Lambda_s M$ in the domain $\mathcal{U}$ of an admissible chart, where

$$\omega|_{\mathcal{U}} = \sum_{j_1 < \ldots < j_s} \omega_{j_1 \ldots j_s}(dx^{j_1} \wedge \ldots \wedge dx^{j_s});$$

we have

$$\begin{aligned}(d\omega)|_{\mathcal{U}} &= d(\omega|_{\mathcal{U}}) \\ &= \sum_{j_1 < \ldots < j_s} \big(d\omega_{j_1 \ldots j_s} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_s} + \omega_{j_1 \ldots j_s} d(dx^{j_1} \wedge \ldots \wedge dx^{j_s})\big) \\ &= \sum_{j_1 < \ldots < j_s} d\omega_{j_1 \ldots j_s} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_s} \,.\end{aligned}$$

With the same remark, we now evaluate d on $d\omega$ in $\mathcal{U}$. We have

$$(d^2\omega)|_{\mathcal{U}} = (dd\omega)|_{\mathcal{U}} = d(d\omega|_{\mathcal{U}}) = \sum_{j_1 < \ldots < j_s} d(d\omega_{j_1 \ldots j_s} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_s}) = 0.$$

So, for any $\omega \in \Lambda_s M$, $d^2\omega = 0$. ■

5.3.4 **Property.** *The exterior derivative commutes with pull-back, i.e., if $\Phi : M \to N$ is a smooth mapping, then*

$$\Phi^* d = d\Phi^*.$$

Proof. (i) Let $\mathcal{V}$ be an open submanifold of N and $\varphi \in C^\infty(\mathcal{V})$. At any point x of open submanifold $\mathcal{U} := \Phi^{-1}(\mathcal{V}) \subset M$, we have

$$(\Phi^* d\varphi)_x = d_{\Phi(x)}\varphi \circ T_x\Phi = d_x(\varphi \circ \Phi|_{\mathcal{U}}) = d_x(\Phi^*\varphi).$$

Then on $\mathcal{U}$,

$$\Phi^* d\varphi = d\Phi^*\varphi$$

and, consequently,

$$d(\Phi^* d\varphi) = d(d\Phi^*\varphi) = 0.$$

(ii) Now let $\mathcal{V}$ be the domain of an admissible chart $\eta = (y^1, \ldots, y^n)$ on N, where, for any $\omega \in \Lambda_s N$,

$$\omega|_{\mathcal{V}} = \sum_{j_1<\ldots<j_s} \omega_{j_1\ldots j_s}(dy^{j_1} \wedge \ldots \wedge dy^{j_s})$$

and

$$(d\omega)|_{\mathcal{V}} = d(\omega|_{\mathcal{V}}) = \sum_{j_1<\ldots<j_s} d\omega_{j_1\ldots j_s} \wedge (dy^{j_1} \wedge \ldots \wedge dy^{j_s}).$$

On $\mathcal{U}$, we have

$$(\Phi^*\omega)|_{\mathcal{U}} = \Phi^*(\omega|_{\mathcal{V}}) = \sum_{j_1<\ldots<j_s} \Phi^*\omega_{j_1\ldots j_s}(\Phi^* dy^{j_1} \wedge \ldots \wedge \Phi^* dy^{j_s})$$

and

$$(\Phi^* d\omega)|_{\mathcal{U}} = \Phi^*(d\omega|_{\mathcal{V}}) = \sum_{j_1<\ldots<j_s} \Phi^* d\omega_{j_1\ldots j_s} \wedge (\Phi^* dy^{j_1} \wedge \ldots \wedge \Phi^* dy^{j_s})$$

Owing to (i), we also have

$$\begin{aligned}
(d\Phi^*\omega)|_{\mathcal{U}} &= d(\Phi^*\omega)|_{\mathcal{U}} \\
&= \sum_{j_1<\ldots<j_s} \big(d\Phi^*\omega_{j_1\ldots j_s} \wedge (\Phi^* dy^{j_1} \wedge \ldots \wedge \Phi^* dy^{j_s}) + \Phi^*\omega_{j_1\ldots j_s} d(\Phi^* dy^{j_1} \wedge \ldots \wedge \Phi^* dy^{j_s})\big) \\
&= \sum_{j_1<\ldots<j_s} \Phi^* d\omega_{j_1\ldots j_s} \wedge (\Phi^* dy^{j_1} \wedge \ldots \wedge \Phi^* dy^{j_s})
\end{aligned}$$

So

$$(\Phi^* d\omega)|_{\mathcal{U}} = (d\Phi^*\omega)|_{\mathcal{U}}.$$

As $\{\mathcal{U} = \Phi^{-1}(\mathcal{V}) \,|\, \mathcal{V}\,\text{coordinate domain on} N\}$ is a covering of M,

$$\Phi^* d\omega = d\Phi^*\omega. \qquad \blacksquare$$

The theory of deRham is the study of *closed* and *exact* forms. Any $\omega \in \Lambda M$ such that $d\omega = 0$, is said to be a closed form. Any $\omega \in \Lambda M$ such that $\omega = d\tau$ for some $\tau \in \Lambda M$, is said to be an exact form. More generally, if for each $x \in M$ there exists an open neighboourhood $\mathcal{U}$ of x where $\omega|_{\mathcal{U}} = d\tau_{\mathcal{U}}$ for some $\tau_{\mathcal{U}} \in \Lambda\mathcal{U}$, ω is said to be a *locally exact* form.

5.3.5 **Theorem.** *Let $\omega \in \Lambda M$. Then ω is locally exact iff it is closed.*

Proof. We will only prove the theorem in $\Lambda_+ M$ (union of all the $\Lambda_s M$'s with $s > 0$).
(i) Let ω be a locally exact form. From $\omega|_{\mathcal{U}} = d\tau_{\mathcal{U}}$, we draw

$$(d\omega)|_{\mathcal{U}} = d(\omega|_{\mathcal{U}}) = d(d\tau_{\mathcal{U}}) = 0$$

and then $d\omega = 0$.
(ii) Let ω be a closed form. For any $x \in M$, let $\mathcal{U}$ be the domain of a spherical chart at x, i.e., $B_o := \xi(\mathcal{U})$ is an open ball with center in the origin of $\mathbb{R}^m$ [41]. In order to show that ω is locally exact, it will suffice to prove that any closed form $\omega_o \in \Lambda_+ B_o$ is exact. Indeed, if we put $\omega_o := (\xi^{-1})^* \omega|_{\mathcal{U}} \in \Lambda B_o$, then

$$d\omega_o := (\xi^{-1})^* d(\omega|_{\mathcal{U}}) = (\xi^{-1})^* (d\omega)|_{\mathcal{U}} = 0,$$

whence

$$\omega_o = d\tau_o$$

for some $\tau_o \in \Lambda B_o$, and finally

$$\omega|_{\mathcal{U}} = \xi^* \omega_o = d(\xi^* \tau_o) = d\tau_{\mathcal{U}}$$

with

$$\tau_{\mathcal{U}} := \xi^* \tau_o \in \Lambda \mathcal{U}.$$

In order to prove that any closed form $\omega_o \in \Lambda_+ B_o$ is exact, one has to use the following

5.3.6 **Poincaré Lemma.** [42] *There exists a mapping*

$$h : \Lambda_+ B_o \to \Lambda B_o$$

such that

$$hd + dh = id\,.$$

Indeed, if

$$d\omega_o = 0,$$

then

$$\omega_o = (hd + dh)\omega_o = dh\omega_o = d\tau_o$$

with

$$\tau_o = h\omega_o\,.$$

∎

Let $X \in \chi(M)$. The *interior product* i_X is another classical derivation (of degree -1) of Cartan calculus, whose action, on any $\omega \in \Lambda_s M$ with $s > 0$, is given by

$$\begin{aligned} i_X \omega(X_2, \ldots, X_s) &= \frac{1}{(s-1)!} \sum_\sigma (sign\sigma)\, \omega(X, X_{\sigma(2)}, \ldots, X_{\sigma(s)}) \\ &= \omega(X, X_2, \ldots, X_s) \end{aligned}$$

---

[41] See footnote [4].

[42] For the proof, see Y. Choquet-Bruhat and C. De Witt-Morette, Analysis, Manifolds and Physics, Part I, (1982) p.224.

that is

$$i_X\omega = \omega(\underbrace{X, \cdot, \ldots, \cdot}_{s})$$

The components of the$(s-1)$-form $i_X\omega$ in the domain $\mathcal{U}$ of an admissible chart ξ, consequently are

$$\begin{aligned}
(i_X\omega)_{j_2,\ldots,j_s} &= (i_X\omega)|_{\mathcal{U}}\left(\frac{\partial}{\partial x^{j_2}}, \ldots, \frac{\partial}{\partial x^{j_s}}\right) = i_X\omega|_{\mathcal{U}}\left(\frac{\partial}{\partial x^{j_2}}, \ldots, \frac{\partial}{\partial x^{j_s}}\right) \\
&= \omega|_{\mathcal{U}}\left(X|_{\mathcal{U}}, \frac{\partial}{\partial x^{j_2}}, \ldots, \frac{\partial}{\partial x^{j_s}}\right) = \omega|_{\mathcal{U}}\left(X^{j_1}\frac{\partial}{\partial x^{j_1}}, \frac{\partial}{\partial x^{j_2}}, \ldots, \frac{\partial}{\partial x^{j_s}}\right) \\
&= X^{j_1}\omega|_{\mathcal{U}}\left(\frac{\partial}{\partial x^{j_1}}, \frac{\partial}{\partial x^{j_2}}, \ldots, \frac{\partial}{\partial x^{j_s}}\right)
\end{aligned}$$

that is

$$(i_X\omega)_{j_2,\ldots,j_s} = X^{j_1}\omega_{j_1,j_2,\ldots,j_s}.$$

Also notice that, for any $f \in C^\infty(M)$,

$$i_X df = df X = Xf.$$

The Lie derivation d_X is the last classical derivation (of degree zero) of Cartan calculus, given by

$$d_X = i_X d + d i_X.$$

Its action, e.g. on $\omega \in \Lambda_1 M$, is given (for any $Y \in \chi M$) by

$$\begin{aligned}
(d_X\omega)Y &= (i_X d\omega)Y + (d i_X\omega)Y = d\omega(X,Y) + \big(d(\omega X)\big)Y \\
&= X(\omega Y) - Y(\omega X) - \omega[X,Y] + Y(\omega X) \\
&= X(\omega Y) - \omega[X,Y]
\end{aligned}$$

The components of 1-form $d_X\omega$ in the domain $\mathcal{U}$ of an admissible chart ξ, consequently are

$$\begin{aligned}
(d_X\omega)_i &= (d_X\omega)|_{\mathcal{U}}\frac{\partial}{\partial x^i} = (d_X\omega|_{\mathcal{U}})\frac{\partial}{\partial x^i} = X|_{\mathcal{U}}\left(\omega|_{\mathcal{U}}\frac{\partial}{\partial x^i}\right) - \omega|_{\mathcal{U}}\left[X|_{\mathcal{U}}, \frac{\partial}{\partial x^i}\right] \\
&= \left(X^j\frac{\partial}{\partial x^j}\right)\omega_i - \omega_j\left[X^h\frac{\partial}{\partial x^h}, \frac{\partial}{\partial x^i}\right]^j \\
&= X^j\frac{\partial\omega_i}{\partial x^j} + \frac{\partial X^j}{\partial x^i}\omega_j\,.
\end{aligned}$$

Also notice that, for any $f \in C^\infty(M)$,

$$d_X f = i_X df = Xf$$

is the Lie derivative of f along X, already introduced in 4.2.
A useful identity is the following

5.3.7 **Property.** *For any* $X, Y \in \chi(M)$, *it is*

$$i_{[X,Y]} = i_X d_Y - d_Y i_X \, .$$

Proof. Owing to the local character of derivations, it will suffice to check that the left and right hand sides take the same value on both f and df (for each $f \in C^\infty(M)$). To this purpose, we have

$$(i_X d_Y - d_Y i_X) f = i_X (d_Y f) - d_Y (i_X f) = 0$$

since both f and $d_Y f$ are in $C^\infty(M)$ and then

$$i_{[X,Y]} f = (i_X d_Y - d_Y i_X) f.$$

Moreover

$$\begin{aligned}
(i_X d_Y - d_Y i_X) df &= i_X (d_Y df) - d_Y (i_X df) = i_X (i_Y d + d i_Y) df - d_Y (d_X f) = i_X d i_Y df - d_Y (X f) \\
&= i_X d(Y f) - Y(X f) = X(Y f) - Y(X f) = [X, Y] f \\
&= i_{[X,Y]} df.
\end{aligned}$$

■

II.

Hamiltonian systems

1 Dynamical systems

1.1 Dynamical systems

Let M be an m-dimensional smooth manifold and

$$c : I \longrightarrow M$$

any motion in M, defined in an open *time interval* $I \subset \mathbb{R}$. The *tangent lift*

$$\dot{c} : I \longrightarrow TM \ : \ t \mapsto \dot{c}(t) \in T_{c(t)}M$$

is a section of tangent bundle $\tau_M : TM \to M$ over c, i.e., $\tau_M \circ \dot{c} = c$. The composition

$$X \circ c : I \longrightarrow TM \ : \ t \mapsto X_{c(t)} \in T_{c(t)}M$$

is another section of τ_M over c, $\tau_M \circ (X \circ c) = c$. If

$$\dot{c} = X \circ c \tag{14}$$

then c is called an *integral curve* of X in M. It is called a *maximal* integral curve if there does not exist any integral curve $k : J \to M$, such that $I \subsetneq J$ and $k|_I = c$. We will call equation (14) a *first-order differential equation* on M, associated with X. From a dynamical point of view, equation (14) expresses the problem of searching for the motions in M whose velocity (at each time) equals the value prescribed by X. Therefore, it will be read as the *equation of the motion of dynamical system*

$$\mathcal{D} := (M, X).$$

The manifold M is called the *phase space* and X the (*velocity*) *field* of $\mathcal{D}$. Any (maximal) solution of the equation of the motion is called a (maximal) *motion* and its image a (maximal) *orbit* of $\mathcal{D}$. The collection of all the maximal orbits is called the *phase portrait* of $\mathcal{D}$.

1.1.1 **Lemma.** *Let $\mathcal{D} = (M, X)$ and $\mathcal{E} = (N, Y)$ be dynamical systems, whose fields are related to each other by a smooth mapping $\Phi : M \to N$ (i.e., $T\Phi \circ X = Y \circ \Phi$). If $c : I \to M$ is a motion of $\mathcal{D}$, then $k := \Phi \circ c : I \to N$ is a motion of $\mathcal{E}$.*

Proof. Owing to hypothesis $\dot{c} = X \circ c$, from $k = \Phi \circ c$ we draw

$$\dot{k} = T\Phi \circ \dot{c} = T\Phi \circ X \circ c = Y \circ \Phi \circ c = Y \circ k.$$ ∎

The above lemma (with its possible enrichments for special Φ's) will often prove to be useful in the study of the motions of a given dynamical system. A first example is the following.
Let $\mathcal{D} = (M, X)$ be a dynamical system. Consider an open submanifold W of M and denote by $X_W \in \chi(W)$ the vector field on W defined by $Tj \circ X_W = X \circ j$ with $j : W \hookrightarrow M$ [(43)]. Then consider the *restricted system* $\mathcal{D}_W = (W, X_W)$, j-related to $\mathcal{D}$

(43) For any $x \in W$, $(X_W)_x = X_x$ in the canonical isomorphism $T_xW \overset{T_xj}{=} T_xM$ (see I. Remark 2.1.7). As to the smoothness of X_W, i.e., of its components in admissible charts, see the next Coordinate expression.

1.1.2 Proposition.
(i) (The image under j of) every motion of $\mathcal{D}_W$ is a motion of $\mathcal{D}$.
(ii) Every motion of $\mathcal{D}$ whose orbit meets W, is a prolongation of (the image under j of) a motion of $\mathcal{D}_W$.

Proof. (i) The statement is a direct consequence of the above lemma.
(ii) Let $c : I \to M$ be a motion of $\mathcal{D}$ such that $c(t_0) \in W$ fro some $t_0 \in I$. By continuity, $c(J) \subset W$ for a suitable $J \subset I$ with $t_0 \in J$. As a consequence, we can consider the induced motion $\gamma : J \to W$. From $c|_J = j \circ \gamma$, we draw

$$(c|_J)^{\cdot} = Tj \circ \dot{\gamma}.$$

From $\dot{c} = X \circ c$, we draw

$$\dot{c}|_J = X \circ c|_J = X \circ j \circ \gamma = Tj \circ X_W \circ \gamma.$$

Then, as $(c|_J)^{\cdot} = \dot{c}|_J$ and Tj is an isomorphism at any point of W, we have

$$\dot{\gamma} = X_W \circ \gamma$$

i.e., γ (which, from a set-theoretical point of view, does not differ from $c|_J$) is a motion of $\mathcal{D}_W$. ∎

The above proposition implies that a local analysis of $\mathcal{D}$, i.e., the study of its motions within an open submanifold W, reduces to the study of restricted system $\mathcal{D}_W$. Typical case of local analysis is the following

1.1.3 Coordinate expression. (44)
Let $\mathcal{U}$ be the domain of an admissible chart $\xi = (x^i)_{i=1\dots,m}$ in M and $j : \mathcal{U} \hookrightarrow M$. Let $\gamma : J \to \mathcal{U}$ be a motion in $\mathcal{U}$, with coordinate expression in ξ given by

$$\xi \circ \gamma = (x^1 \circ \gamma, \dots, x^m \circ \gamma) =: (\gamma^1, \dots, \gamma^m).$$

From the decompositions

$$\dot{\gamma} = \frac{d\gamma^i}{du}\left(\frac{\partial}{\partial x^i} \circ \gamma\right)$$
$$X_{\mathcal{U}} = X^i \frac{\partial}{\partial x^i}$$

we draw

$$X_{\mathcal{U}} \circ \gamma = (X^i \circ \gamma)\left(\frac{\partial}{\partial x^i} \circ \gamma\right) = \left(X^i_\xi \circ (\gamma^1, \dots, \gamma^m)\right)\left(\frac{\partial}{\partial x^i} \circ \gamma\right)$$

where $(X^i_\xi := X^i \circ \xi^{-1})$ are the coordinate expression of components (X^i) of X in ξ. Then γ is a motion of $\mathcal{D}_{\mathcal{U}}$, if, and only if,

$$\frac{d\gamma^i}{du} = X^i_\xi \circ (\gamma^1, \dots, \gamma^m). \tag{15}$$

This shows that the coordinate local analysis of a dynamical system $\mathcal{D}$, reduces to the study of a system (15) of m ordinary, first-order, normal, differential equations in m unknown real-valued functions $(\gamma^1, \dots, \gamma^m)$. ∎

(44) See also the coordinate expression of a smooth curve in I 2.3.1 and the local decomposition law of a vector field in I 4.2.

1.1.4 **Note.** If $x \in M$ is a *non-singular* point, i.e. $X_x \neq 0$, then there exists an admissible chart $(\mathcal{U}, \xi)$ on M at x, such that, $X_{\mathcal{U}}$ is one of the coordinate vector fields, say

$$X_{\mathcal{U}} = \frac{\partial}{\partial x^1}$$

(straightening theorem [45]). The equation of the motion of $\mathcal{D}_{\mathcal{U}} = \left(\mathcal{U}, \frac{\partial}{\partial x^1}\right)$ is 'integrable by quadratures', for its expression in ξ is

$$\frac{d\gamma^i}{dt} = \delta_1^i \qquad (i = 1, \ldots, m),$$

and then its solutions correspond in $\xi(\mathcal{U})$ to motions which take place, with velocity $\frac{\partial}{\partial u^1}$, on straight lines of direction $\delta_1 = (\delta_1^i)$. ■

1.2 Determinism

Let $\mathcal{D} = (M, X)$ be a dynamical system and $(t_0, x) \in \mathbb{R} \times M$ arbitrarily chosen *Cauchy data.* Any motion $c : I \to M$, starting at time $t_0 \in I$ from $c(t_0) = x$, and (maximal) solution of differential equation

$$\dot{c} = X \circ c\,,$$

will be said to be a (maximal) solution of *Cauchy problem* (M, X, t_0, x). We shall conventionally choose $t_0 = 0$ as initial instant, and we shall denote a Cauchy problem by (M, X, x). From the local analysis of a dynamical system, we draw the following

1.2.1 **Local determinism theorem.** *For each $x \in M$, there exist solutions of (M, X, x). Any two of them locally agree* [46].

Proof. This is basically a result of the theory of ordinary differential equations. On the domain $\mathcal{U}$ of an admissible chart ξ at x, consider Cauchy problem $(\mathcal{U}, X_{\mathcal{U}}, x)$, i.e.,

$$\begin{cases} \dot{\gamma} = X_{\mathcal{U}} \circ \gamma \\ \gamma(0) = x \end{cases} \tag{16}$$

whose coordinate expression in ξ is [47]

$$\begin{cases} \dfrac{d\gamma^i}{dt} = X_\xi^i \circ (\gamma^1, \ldots, \gamma^m) \\ \gamma^i(0) = x^i(x). \end{cases} \tag{17}$$

From the theory of ordinary differential equations, it is known that problem (17) admits solutions in $\xi(\mathcal{U})$ and that any two of them

$$(\gamma_\alpha^1, \ldots, \gamma_\alpha^m) : I_\alpha \longrightarrow \xi(\mathcal{U}) \qquad (\alpha = 1, 2)$$

locally agree. Consequently, curves

$$\gamma_\alpha := \xi^{-1} \circ (\gamma_\alpha^1, \ldots, \gamma_\alpha^m) : I_\alpha \longrightarrow \mathcal{U} \qquad (\alpha = 1, 2)$$

are solutions to problem (11) and locally agree. Hence, since solutions of $(\mathcal{U}, X_{\mathcal{U}}, x)$ are solutions of (M, X, x) too [48], the statement follows. ■

---

(45) See Brickell and R.S.Clark, Differential Manifolds, (1970), p.140.

(46) This amounts to saing that, if $c_1 : I_1 \to M$ and $c_2 : I_2 \to M$ are solutions of (M, X, x), then there exists an open interval I_0, with $0 \in I_0 \subset I_1 \cap I_2$, where $c_1|_{I_0} = c_2|_{I_0}$.

(47) See II 1.1.3.

(48) See II 1.1.2(i).

1.2.2 **Lemma.** *Let $c : I \to M$ be a smooth curve through $x_0 := c(t_0)$, with $t_0 \neq 0$. By means of the translation*

$$\tau_{t_0} : \mathbb{R} \longrightarrow \mathbb{R} \; : t \mapsto t + t_0$$

which maps the open interval

$$I - t_0 := \tau_{t_0}^{-1}(I) = \{t \in \mathbb{R} \,|\, t + t_0 \in I\}$$

onto I, define a re-parametrization of c,

$$c_0 := c \circ \tau_{t_0} : I - t_0 \longrightarrow M \; : \; t \mapsto c(t + t_0),$$

with $c_0(0) = x_0$. The curve c is a solution of (M, X, t_0, x_0) iff c_0 is a solution of (M, X, x_0).

Proof. From $c_0 := c \circ \tau_{t_0}$, it follows that

$$\dot{c}_0 := Tc \circ \dot{\tau}_{t_0} = Tc \circ \frac{d}{dt} \circ \tau_{t_0} = \dot{c} \circ \tau_{t_0} \, .$$

So $\dot{c} = X \circ c$ iff $\dot{c} \circ \tau_{t_0} = X \circ c \circ \tau_{t_0}$ iff $\dot{c}_0 = X \circ c_0$. ∎

The intermediate passage is the following.

1.2.3 **Proposition.** *If $c_1 : I_1 \to M$ and $c_2 : I_2 \to M$ are solutions of Cauchy problem (M, X, x), then they agree on the whole intersection $I_1 \cap I_2$:*

$$c_1|_{I_1 \cap I_2} = c_2|_{I_1 \cap I_2}.$$

Proof. Let

$$I_0 := \{t \in I_1 \cap I_2 \,|\, c_1(t) = c_2(t)\}.$$

Notice that $0 \in I_0 \subset I_1 \cap I_2$ and that $I_1 \cap I_2$, being an open interval, is a connected space. So one can prove the statement $I_0 = I_1 \cap I_2$ by showing that I_0 is both a closed and an open subset of $I_1 \cap I_2$. Define a smooth mapping $\lambda : I_1 \cap I_2 \to M \times M$ by putting, for any $t \in I_1 \cap I_2$, $\lambda(t) = \big(c_1(t), c_2(t)\big)$. Clearly, we have $I_0 = \lambda^{-1}(\Delta)$, where $\Delta := \{(x, x) \in M \times M\}$ is the diagonal of $M \times M$. The diagonal Δ is a closed subset of $M \times M$, for M is a Hausdorff space. As a consequence, its inverse image I_0 by continous mapping λ is a closed subset of $I_1 \cap I_2$. Let now $t_0 \in I_0$. Consider the re-parametrization c_{1_0}, c_{2_0} of c_1, c_2 as in the previous lemma. The same lemma ensures that c_{1_0} and c_{2_0} are both solutions of Cauchy problem (M, X, x_0) with $x_0 := c_1(t_0) = c_2(t_0)$. Owing to local determinism theorem, there exists a suitably small open interval $(-\epsilon, \epsilon)$ such that., $c_{1_0}(t) = c_{2_0}(t)$ for all $t \in (-\epsilon, \epsilon)$, i.e., $c_1(t+t_0) = c_2(t+t_0)$ for all $t+t_0 \in (t_0-\epsilon, t_0+\epsilon)$ and then $(t_0-\epsilon, t_0+\epsilon) \subset I_0$. So I_0, as union of open intervals, is an open subset of $I_1 \cap I_2$. ∎

1.2.4 **Global Determinism Theorem.** *For each $x \in M$, there exists a unique maximal motion of $\mathcal{D}$ starting from x.*

Proof. Let $\{c_\alpha : I_\alpha \to M\}$ be the collection of all the solutions of (M, X, x),

$$\begin{cases} \dot{c}_\alpha = X \circ c_\alpha \\ c_\alpha(0) = x. \end{cases}$$

Put

$$I_x := \cup_\alpha I_\alpha$$

and note that I_x is an open interval containing zero. Owing to the previous proposition, all of the c_α's are restrictions of the smooth curve

$$\Phi_x : I_x \to M$$

defined by

$$\Phi_x|_{I_\alpha} := c_\alpha \,.$$

First Φ_x is a solution of (M, X, x), since for all $t \in I_x$ we have

$$\dot{\Phi}_x(t) = \dot{c}_\alpha(t) = X \circ c_\alpha(t) = X \circ \Phi_x(t)$$

and

$$\Phi_x(0) = c_\alpha(0) = x.$$

Moreover Φ_x is a maximal solution, for any solution is one of its restrictions and, finally, Φ_x is the unique maximal solution, for the same reason. ∎

A consequence of global determinism is the following globalization of II 1.2.1, in the case of a covering map.

1.2.5 **Corollary.** *Let $\mathcal{D} = (M, X)$ and $\mathcal{E} = (N, Y)$ be dynamical systems, related to each other by a covering map $h : M \to N$. If, for any $y \in N$ and $x \in h^{-1}(y)$, $\Psi_y : J_y \to N$ and $\Phi_x : I_x \to M$ are the maximal solutions of $\mathcal{E}$ and $\mathcal{D}$ starting from y and x respectively, then*

$$\Psi_y = h \circ \Phi_x \,.$$

Proof. On the one hand, $h \circ \Phi_x : I_x \to N$ is a motion of $\mathcal{E}$ (owing to Lemma 1.2.1) and starts from $h \circ \Phi_x(0) = h(x) = y$. Global determinism then ensures that it is a restriction of Ψ_y, i.e.,

$$I_x \subset J_y \tag{18}$$

and

$$h \circ \Phi_x = \Psi_y|_{I_x} \,. \tag{19}$$

On the other hand, the lift theorem [49] ensures that there exixts a unique lift $c : J_y \longrightarrow M$ of Ψ_y by h,

$$h \circ c = \Psi_y \,, \tag{20}$$

starting from $c(0) = x$. Notice that c is a motion of $\mathcal{D}$, for the time derivative of (20) yields

$$Th \circ \dot{c} = \dot{\Psi}_y = Y \circ \Psi_y = Y \circ h \circ c = Th \circ X \circ c$$

whence, Th being an isomorphism at any point of M, $\dot{c} = X \circ c$. If c is a motion of $\mathcal{D}$ starting from x, it must be a restriction of Φ_x, and then

$$J_y \subset I_x \,. \tag{21}$$

Owing to (18) and (21), one has

$$I_x = J_y \,,$$

and then equality (19) reads

$$h \circ \Phi_x = \Psi_y \,.$$ ∎

[49] See I 3.3.6.

1.2.6 **Lemma.** *Let* $\Phi_x : I_x \to M$ *be the maximal solution of* (M, X, x)*. For any*

$$x_0 = \Phi_x(t_0)$$

(with $0 \neq t_0 \in I_x$*), the maximal solution of* (M, X, x_0) *is given by*

$$\Phi_{x_0} = \Phi_x \circ \tau_{t_0} \,.$$

Proof. Put

$$c_0 := \Phi_x \circ \tau_0 : I_x - t_0 \longrightarrow M.$$

Owing to Lemma 1.2.2, c_0 is a solution of (M, X, x_0). Consequently, it is a restriction of Φ_{x_0}, i.e.,

$$I_x - t_0 \subset I_{x_0} \tag{22}$$

and

$$c_0 = \Phi_{x_0}|_{I_x - t_0} \,. \tag{23}$$

Now, evaluating (23) in $-t_0 \in I_x - t_0$, we have

$$\Phi_{x_0}(-t_0) = c_0(-t_0) = \Phi_x(-t_0 + t_0) = \Phi_x(0) = x.$$

As above, $x = \Phi_{x_0}(-t_0)$ implies $I_x \supset I_{x_0} + t_0$, i.e.,

$$I_x - t_0 \supset I_{x_0} \,. \tag{24}$$

Owing to (22) and (24), we have $I_{x_0} = I_x - t_0$ and then (23) reads

$$\Phi_{x_0} = c_0 \,.$$

■

1.2.7 **Remark.** It is useful to explicitly remark that, during the proof of the above lemma, we have obtained the following 'time reversibility' law

$$x_0 = \Phi_x(t_0) \iff x = \Phi_{x_0}(-t_0).$$

■

With the aid of the above lemma, we shall now show that, in the phase space of $\mathcal{D}$, maximal orbits are separated.

1.2.8 **Orbital Determinism Theorem.** *For each* $x \in M$*, there exists a unique maximal orbit of* $\mathcal{D}$ *containing* x*.*

Proof. On one hand,

$$x \in \Gamma := \Phi_x(I_x)$$

for $x = \Phi_x(0)$. On the other hand, if

$$x \in \Gamma_0 := \Phi_{x_0}(I_{x_0}) \,,$$

then $x = \Phi_{x_0}(-t_0)$, for some $-t_0 \in I_{x_0}$, or equivalently $x_0 = \Phi_x(t_0)$, whence owing to Lemma 1.2.6, $I_{x_0} = I_x - t_0$ and $\Phi_{x_0} = \Phi_x \circ \tau_{t_0}$. As a consequence

$$\Gamma_0 = \Phi_{x_0}(I_{x_0}) = \Phi_x\big(\tau_{t_0}(I_x - t_0)\big) = \Phi_x(I_x) = \Gamma.$$

■

1.3 Flow of a dynamical system

We shall now collect all the maximal motions $\{\Phi_x\,;\, x \in M\}$ of a dynamical system $\mathcal{D} = (M, X)$ into one mapping

$$\Phi : D \longrightarrow M$$

defined on

$$D := \bigcup_{x \in M} (I_x \times \{x\})$$

by putting, for any $(t, x) \in D$,

$$\Phi(t, x) := \Phi_x(t).$$

The mapping Φ is called the *flow* of $\mathcal{D}$. Firstly, we have the following

1.3.1 **Theorem.** *The set D is an open submanifold of $\mathbb{R} \times M$, and Φ is a smooth mapping.*

Proof. We shall only sketch the proof, whose spirit is to show that, for any $(t, x) \in D$,
(o) there exists an open interval $I \ni t$ and an open neighbourhood $W \ni x$, such that $I \times W \subset D$ and $\Phi|_{I \times W}$ is smooth.
(i) First, one proves that the above condition holds true at any $(0, x) \in D$ (this is basically a result from the theory of ordinary differential equations, taken back to M by an admissible chart at x).
(ii) Then one proves that, for any $x \in M$, non-void subset

$$K_x := \{t \in I_x \,|\, \text{condition (o) is satisfied at } (t, x)\}$$

is both (obviously) open and closed in I_x (the latter result follows [50] from (i) above and a forthcoming pseudo-group property of Φ) and then, owing to the connectedness of I_x, $K_x = I_x$. This shows that condition (o) holds true at any $(t, x) \in D$. ∎

1.3.2 **Remark.** It is useful to explicitly remark that, owing to conditon (o), for any $x \in M$ and $t \in I_x$, there exixts a real number $\epsilon > 0$ and an open neighbourhood W of x such that, for all $x \in W$,

$$(t - \epsilon, t + \epsilon) \subset I_x\,.$$ ∎

For any $t \in \mathbb{R}$, define

$$\Phi_t : D_t \to M$$

on

$$D_t := \{x \in M \,|\, t \in I_x\}$$

by putting, for any $x \in D_t$,

$$\Phi_t(x) := \Phi(t, x) = \Phi_x(t).$$

1.3.3 **Corollary.** *D_t is an open submanifold of M, and Φ_t is a smooth mapping.*

Proof. First notice that the above Remark 1.3.2 implies that, for any $x \in D_t$, there exists an open neighbourhood W of x such that $t \in I_x$, for all $x \in W$, i.e., $W \subset D_t$. This shows that D_t is open in M. Moreover

$$\Phi_t = \Phi \circ \varphi_t|_{D_t}$$

where

$$\varphi_t : M \longrightarrow \mathbb{R} \times M \;:\; x \mapsto \varphi_t(x) = (t, x).$$

So Φ_t, composition of smooth mappings, is smooth. ∎

[50] See S.Lang, Differential Manifolds, (1972), p.86.

1.3.4 **Pseudo group property.**
(i) $\Phi_0 = id_M$
(ii) For any $t_1, t_2 \in \mathbb{R}$, *on* $D_{t_1+t_2} \cap D_{t_2}$, *it is*

$$\Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2} .$$

Proof. (i) Trivial.
(ii) Let us first check that

$$\Phi_{t_2}(D_{t_1+t_2} \cap D_{t_2}) \subset D_{t_1} .$$

To this end, let $x \in D_{t_2} \cap D_{t_1+t_2}$, i.e., $t_2, t_1 + t_2 \in I_x$. Then put

$$x_2 := \Phi_{t_2}(x) = \Phi_x(t_2)$$

and note that

$$t_1 \in I_x - t_2 = I_{x_2}$$

whence

$$x_2 \in D_{t_1} .$$

So both the left and the right hand side of (ii) are defined on $D_{t_1+t_2} \cap D_{t_2}$ (if non-empty), where they agree, for

$$\begin{aligned}\Phi_{t_1}\big(\Phi_{t_2}(x)\big) &= \Phi_{t_1}(x_2) = \Phi_{x_2}(t_1) = \Phi_x(t_1 + t_2) \\ &= \Phi_{t_1+t_2}(x).\end{aligned}$$

■

1.3.5 **Corollary.** *For any* $t \in \mathbb{R}$, Φ_t *induces a diffeomorphism of* D_t *onto* D_{-t}.

Proof. Let us first check that $\Phi_t(D_t) \subset D_{-t}$. To this end, let $x \in D_t$, i.e., $t \in I_x$. Then put $x_t := \Phi_t(x) = \Phi_x(t)$ and recall that $I_{x_t} = I_x - t \ni -t$, whence $x_t \in D_{-t}$. Consequently, since D_{-t} is an open and then smoothness preserving submanifold of M, the smooth mapping $\Phi_t : D_t \to M$ induces a smooth mapping $\Phi_t : D_t \to D_{-t}$. Now let us notice that, owing to Theorem 1.3.4, such induced mappings satisfy

$$\begin{aligned}\Phi_{-t} \circ \Phi_t &= \Phi_0|_{D_t} = \mathrm{id}_{D_t} , \\ \Phi_t \circ \Phi_{-t} &= \Phi_0|_{D_{-t}} = \mathrm{id}_{D_{-t}} .\end{aligned}$$

Then $\Phi_t : D_t \to D_{-t}$ is a smooth mapping which admits of a smooth inverse $\Phi_{-t} : D_{-t} \to D_t$, and then is a diffeomorphism. ■

Any diffeomorphism between open subsets of M is called a *local transformation in* M. So Theorem 1.3.4 and Corollary 1.3.5 can be rephrased by saying that Φ defines a one-parameter pseudogroup $\{\Phi_t; t \in \mathbb{R}\}$ of local transformations in M. A special case is the one of a *complete* vector field X on M (or complete dynamical system), characterized by

$$I_x = \mathbb{R}$$

for all $x \in M$, or equivalently

$$D_t = M$$

for all $t \in \mathbb{R}$, or

$$D = \mathbb{R} \times M.$$

In this case, the flow is a *one-parameter group of transformations of* M, i.e., a smooth mapping

$$\Phi : \mathbb{R} \times M \longrightarrow M$$

such that

$$\Phi_t := \Phi \circ \varphi_t : M \longrightarrow M$$

satisfy

$$\Phi_0 = \mathrm{id}_M$$
$$\Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2} .$$

The family of transformations $\{\Phi_t \,;\, t \in \mathbb{R}\}$ form an Abelian group of diffeomorphisms of M onto itself. The *trajectories*

$$\Phi_x := \Phi \circ \varphi_x : \mathbb{R} \longrightarrow M \quad , \quad x \in M$$

(with $\varphi_x : t \in \mathbb{R} \to (t, x) \in \mathbb{R} \times M$) are the maximal motions of the system.
This result can be inverted. Let Φ be a one-parameter group of transformations of a manifold M. Define the *infinitesimal generator* of Φ as the vector field $X \in \chi(M)$ given, for any $x \in M$, by

$$X_x := \dot{\Phi}_x(0)$$

As to the smoothness of X see the following coordinate expression.

1.3.6 **Proposition.** *The infinitesimal generator X of a one-parameter group Φ of transformations of M, is a complete vector field on M whose flow is Φ itself.*

Proof. We have to prove that, for any $x \in M$,

$$\dot{\Phi}_x = X \circ \Phi_x .$$

To this end, let $t_0 \in \mathbb{R}$, $x_0 := \Phi_x(t_0)$, $\Phi_{x_0} = \Phi_x \circ \tau_{t_0}$, then

$$X \circ \Phi_x(t_0) = X_{x_0} = \dot{\Phi}_{x_0}(0) = \dot{\Phi}_x \circ \tau_{t_0}(0) = \dot{\Phi}_x(t_0) .$$

∎

1.3.7 **Coordinate expression.** Let $\xi = (x^1, \ldots, x^m) : \mathcal{U} \to \mathbb{R}^m$ be an admissible chart on M. For any $x \in \mathcal{U}$, the coordinate expression of Φ_x – restricted to a suitably small open interval $I \ni 0$ so that $\Phi_x(I) \subset \mathcal{U}$ – is

$$\Phi^i_x := x^i \circ \Phi_x|_I = x^i \circ \Phi \circ \varphi_x|_I \quad , \quad (i = 1, \ldots, m).$$

Then

$$X^i_x = \left.\frac{d\Phi^i_x}{dt}\right|_0 = d_{(0,x)}(x^i \circ \Phi)\,(\dot{\varphi}_x(0)) = d_{(0,x)}(x^i \circ \Phi) \left(\left.\frac{\partial}{\partial t}\right|_{(0,x)} \right) = \left.\frac{\partial}{\partial t}(x^i \circ \Phi)\right|_{(0,x)}$$

whence

$$X^i = \frac{\partial}{\partial u}(x^i \circ \Phi) \circ \varphi_0|_{\mathcal{U}} .$$

Notice that the smoothness of components (X^i) ensures that $X \in \chi(M)$. ∎

We will now find an intrinsic link between the action of the flow Φ of a given dynamical system (M, X) on the exterior algebra ΛM and the Lie derivative d_X. Indeed each local transformation

$$\Phi_t : D_t \to M \quad , \quad t \in \mathbb{R}$$

acts on ΛM by pull-back [(51)]

$$\Phi_t^* : \omega \in \Lambda M \mapsto \phi_t^* \omega \in \Lambda D_t \simeq \Lambda M \, .$$

So, if $\omega \in \Lambda_r M$ (in the non trivial case $r \geq 0$) [(52)],

$$(\Phi_t^* \omega)_x = \omega_{\Phi_t(x)} \circ (T_x \Phi_t)^r$$

for any $t \in \mathbb{R}$ and $x \in D_t$ or, equivalently, for any $x \in M$ and $t \in I_x$. Notice that, for any $x \in M$, the mapping

$$t \in I_x \mapsto (\Phi_t^* \omega)_x \in (ST_r^0 M)_x$$

takes its values in the vector subspace $(ST_r^0 M)_x \subset (T_r^0 M)_x$ of skew-symmetric $\binom{0}{r}$-tensors at x. Its (ordinary) derivative at any $t_0 \in I_x$ is

$$\frac{d}{dt} (\Phi_t^* \omega)_x \bigg|_{t=t_0} := \lim_{t \to t_0} \frac{1}{t - t_0} \big((\Phi_t^* \omega)_x - (\Phi_{t_0}^* \omega)_x \big) \in (ST_r^0 M)_x \, .$$

If, in particular, we put

$$(L_X \omega)_x := \frac{d}{dt} (\Phi_t^* \omega)_x \bigg|_{t=0} = \lim_{t \to t_0} \frac{1}{t} \big((\Phi_t^* \omega)_x - \omega_x \big)$$

we have

1.3.8 **Lemma.** *If $x_0 := \Phi_x(t_0)$, for some $t_0 \in I_x$, then*

$$\frac{d}{dt} (\Phi_t^* \omega)_x \bigg|_{t=t_0} = (L_X \omega)_{x_0} \circ (T_x \Phi_{t_0})^r$$

Proof. From $x_0 = \Phi_x(t_0)$, it follows that

$$x = \Phi_{x_0}(-t_0)$$

and then, for any $t \in I_x$,

$$\Phi_x(t) = \Phi_{x_0}(t - t_0)$$

that is

$$\Phi_t(x) = \Phi_{t-t_0}(x_0)$$

and, on $D_t \cap D_{t_0} \ni x$,

$$\Phi_t = \Phi_{t-t_0} \circ \Phi_{t_0}$$

---

(51) See I 5.1.

(52) $(T_x \Phi_t)^r := \underbrace{T_x \Phi_t \times \ldots \times T_x \Phi_t}_{r \text{ times}}$ is absent if $r = 0$.

whence

$$T_x\Phi_t = T_{x_0}\Phi_{t-t_0} \circ T_x\Phi_{t_0} .$$

As a consequence,

$$(\Phi_t^*\omega)_x = \omega_{\Phi_t(x)} \circ (T_x\Phi_t)^r = \omega_{\Phi_{t-t_0}(x_0)} \circ (T_{x_0}\Phi_{t-t_0})^r \circ (T_x\Phi_{t_0})^r = (\Phi_{t-t_0}^*\omega)_{x_0} \circ (T_x\Phi_{t_0})^r$$

and then

$$\frac{d}{dt}(\Phi_t^*\omega)_x\bigg|_{t=t_0} = \frac{d}{dt}(\Phi_{t-t_0}^*\omega)_{x_0}\bigg|_{t-t_0=0} \circ (T_x\Phi_{t_0})^r = (L_X\omega)_{x_0} \circ (T_x\Phi_{t_0})^r$$ ∎

If we put

$$L_X\omega : x \in M \mapsto (L_X\omega)_x \in ST_r^0M$$

we obtain a scalar form $L_X\omega \in \Lambda_rM$.

1.3.9 **Coordinate expression.** Let, e.g., $\omega \in \Lambda_1M$. Let $\xi = (x^i)_{i=1,\dots,m}$ be an admissible chart at x. For suitably small $t \in I_x$, also $\Phi_x(t) = \Phi_t(x) \in \mathcal{U}$ and then $(\Phi_t^*\omega)_x$ has components in ξ, given by [(53)]

$$(\Phi_t^*\omega)_i(x) = \frac{\partial\Phi_t^j}{\partial x^i}\bigg|_x \omega_j(\Phi_t x) = \frac{\partial\Phi^j}{\partial x^i}\bigg|_{(t,x)} \omega_j\big(\Phi(t,x)\big)$$

where $\Phi^j = x^j \circ \Phi|$ ($\Phi|$ stands for a suitable restriction of Φ). Then [(54)]

$$\begin{aligned}(L_X\omega)_i(x) &= \frac{d}{dt}(\Phi_t^*\omega)_i(x)\bigg|_0 = \frac{\partial^2\Phi^j}{\partial t\partial x^i}\bigg|_{(0,x)} \omega_j(x) + \frac{\partial\Phi^j}{\partial x^i}\bigg|_{(0,x)} \frac{\partial\omega_j}{\partial x^h}\bigg|_x \frac{\partial\Phi^h}{\partial t}\bigg|_{(0,x)} \\ &= \frac{\partial X^j}{\partial x^i}\bigg|_x \omega_j(x) + \delta_j^i \frac{\partial\omega_j}{\partial x^h}\bigg|_x X^h(x) \\ &= \left(\frac{\partial X^j}{\partial x^i}\omega_j + \frac{\partial\omega_i}{\partial x^j}X^j\right)(x)\end{aligned}$$

The smoothness of components $(L_X\omega)_i$ ensures that $L_X\omega \in \Lambda_1M$. ∎

Now consider

$$L_X : \omega \in \Lambda M \mapsto L_X\omega \in \Lambda M$$

1.3.10 **Lemma.**

$$L_X = d_X .$$

Proof. Let, e.g., $\omega \in \Lambda_1M$. The statement $L_X = d_X$, simply follows from the above coordinate expression, which shows that, in any admissible chart, $(L_X\omega)_i = (d_X\omega)_i$ [(55)]. ∎

The interest of the above point of view about Lie derivative d_X, stays in the following result. An exterior form $\omega \in \Lambda M$ is said to be Φ-*invariant* if for all $t \in \mathbb{R}$,

$$\Phi_t^*\omega = \omega|_{D_t}$$

(53) See I 5.1.4.
(54) See 1.3.6.
(55) See I 5.3.

that is

$$(\Phi_t^* \omega)_x = \omega_x$$

for all $t \in \mathbb{R}$ and $x \in D_t$ (or for all $x \in M$ and $t \in I_x$), or equivalently

$$\frac{d}{dt}(\Phi_t^* \omega)_x = 0 \quad , \quad \forall x \in M$$

As a consequence, owing to Lemmas 1.3.8 and 1.3.9,

1.3.11 **Theorem.** *An exterior form $\omega \in \Lambda M$ is Φ-invariant if, and only if*

$$d_X \omega = 0. \qquad \blacksquare$$

Now a similar reasoning on the action of Φ on vector fields, will lead us to extend the action of L_X on χM. Let $Y \in \chi(M)$. For any $x \in (M)$, put

$$(\Phi_{-t\,*} Y)_x := (T\Phi_{-t} \circ Y \circ \Phi_t)_x = T\Phi_{-t}(Y_{\Phi_t(x)}).$$

Consider the mapping

$$t \in I_x \mapsto (\Phi_{-t\,*} Y)_x \in T_x M$$

and its ordinary derivative at $t = 0$

$$(L_X Y)_x := \frac{d}{dt}(\Phi_{-t\,*} Y)_x \bigg|_{t=0} = \lim_{t \to 0} \frac{1}{t}\big((\Phi_{-t\,*} Y)_x - Y_x\big) \in T_x M$$

1.3.12 **Lemma.** *If $x_0 := \Phi_x(t_0)$, for some $t_0 \in I_x$, then*

$$\frac{d}{dt}(\Phi_{-t\,*} Y)_x \bigg|_{t=t_0} = T\Phi_{-t_0}(L_X Y)_{x_0}\,.$$

Proof. The same considerations as in the proof of Lemma 1.3.8, yield $\Phi_t(x) = \Phi_{t-t_0}(x_0)$ for any $t \in I_x$, and $\Phi_{-t} = \Phi_{-t_0} \circ \Phi_{-t+t_0}$ on $D_{-t} \cap D_{-t+t_0} \ni \Phi_t(x)$. As a consequence,

$$(\Phi_{-t\,*} Y)_x = T\Phi_{-t_0} \circ T\Phi_{-t+t_0}(Y_{\Phi_{t-t_0}(x_0)}) = T\Phi_{-t_0}(\Phi_{-t+t_0\,*} Y)_{x_0}$$

and then

$$\frac{d}{dt}(\Phi_{-t\,*} Y)_x \bigg|_{t=t_0} = T\Phi_{-t_0}\left(\frac{d}{dt}(\Phi_{-t+t_0\,*} Y)_{x_0} \bigg|_{t-t_0=0}\right) = T\Phi_{-t_0}(L_X Y)_{x_0} \qquad \blacksquare$$

If we now put

$$L_X Y : x \in M \mapsto (L_X Y)_x \in TM$$

we obtain a vector field $L_X Y \in \chi(M)$.

1.3.13 **Coordinate expression.** In a chart ξ as in 1.3.9., $(\Phi_{-t\,*}Y)_x$ has components given by (56)

$$(\Phi_{-t\,*}Y)^i_x = Y^j\big(\Phi_t(x)\big)\frac{\partial \Phi^i}{\partial x^j}\big(-t, \Phi_t(x)\big)$$

and then

$$(L_X Y)^i(x) = -Y^i(x)\left.\frac{\partial^2 \Phi^i}{\partial t \partial x^j}\right|_{(0,x)} + \left.\frac{\partial Y^j}{\partial x^h}\right|_x \left.\frac{\partial \Phi^h}{\partial t}\right|_{(0,x)} \left.\frac{\partial \Phi^i}{\partial x^j}\right|_{(0,x)}$$
$$= \left(-Y^j\frac{\partial X^i}{\partial x^j} + \frac{\partial Y^i}{\partial x^j}X^j\right)(x)$$

The smoothness of components $(L_X Y)^i$ ensures that of $L_X Y \in \chi(M)$. ■

Consider now

$$L_X : Y \in \chi(M) \mapsto L_X Y \in \chi(M)$$

1.3.14 **Lemma.** $L_X = [X, \cdot]$.

Proof. Let $Y \in \chi(M)$. The statement $L_X Y = [X, Y]$, simply follows from the above coordinate expression, which shows that, in any admissible chart, $(L_X Y)^i = [X, Y]^i$ (57). ■

1.3.15 **Remark.** A simple calculation of components, also shows that $L_X : \chi(M) \to \chi(M)$ is a derivation on the $C^\infty(M)$-module $\chi(M)$, i.e., $\mathbb{R}$-linear

$$L_X(a_\alpha Y^\alpha) = a_\alpha(L_X Y^\alpha)$$

(with $a_\alpha \in \mathbb{R}$, $Y^\alpha \in \chi(M)$) and Leibniz

$$L_X(fY) = f(L_X Y) + (L_X f)Y$$

(with $f \in C^\infty(M)$, $Y \in \chi(M)$). ■

The interest of the above view of Lie bracket $[X, \cdot]$ as a Lie derivative on $\chi(M)$, lies in the following result. A vector field $Y \in \chi(M)$ is said to be Φ-*invariant* if for any $t \in \mathbb{R}$,

$$\Phi_{-t\,*}Y = Y|_{D_t}$$

that is, for all $t \in \mathbb{R}$ and $x \in D_t$, or for all $x \in M$ and $\in I_x$,

$$(\Phi_{-t\,*}Y)_x = Y_x$$

or, equivalently,

$$\frac{d}{dt}(\Phi_{-t\,*}Y)_x = 0$$

for all $x \in M$. Owing to Lemmas 1.3.12 and 1.3.14, we have

(56) See I 2.3.4.?
(57) See I 4.4.?

1.3.16 **Theorem.** *A vector filed Y is Φ-invariant iff*

$$[X, Y] = 0.$$ ∎

An interesting consequence of the above theorem, is the following

1.3.17 **Corollary.** *Let $X, Y \in \chi(M)$ be complete vector fields. Their flows Φ, Ψ commute,*

$$\Phi_t \circ \Psi_s = \Psi_s \circ \Phi_t \,, \qquad \forall t, s \in \mathbb{R}$$

if, and only if, X, Y commute,

$$[X, Y] = 0.$$ ∎

The proof can be drawn from the following

1.3.18 **Remark.** Let $Y \in \chi(M)$ be a complete vector field, and $h : M \to M$ a transformation of M. As is known, Y is the infinitesimal generator of a one-parameter group $\{\Psi_s \,, s \in\mathbb{R}\}$ of transformations of M, and then, for any $x \in M$, Y_x is tangent at x to the trajectory

$$s \in\mathbb{R} \mapsto \Psi_s(x) \in M.$$

As a consequence, the *push-forward* of Y by h:

$$h_* Y := Th \circ Y \circ h^{-1} \in \chi(M)$$

turns out to be the infinitesimal generator of the one-parameter group $\{\chi_s \,, s \in\mathbb{R}\}$ given by $\chi_s = h \circ \Psi_s \circ h^{-1}$, since, for any $y \in M$, $(h_* Y)_y = T_x h(Y_y)$ (with $x := h^{-1}(y)$) is tangent at y to the trajectory

$$s \in\mathbb{R} \mapsto h \circ \Psi_s(x) = \chi_s(y) \in M.$$

Therefore, Y is *h-invariant*:

$$h^* Y = Y$$

(i.e., Y is h-related to itself: $Th \circ Y = Y \circ h$), iff

$$\chi_s = \Psi_s \qquad \forall s \in\mathbb{R}$$

(i.e., Ψ_s commutes with h: $h \circ \Psi_s = \Psi_s \circ h$, for any $s \in\mathbb{R}$). ∎

Let $\mathcal{P}$ be the phase portrait of dynamical system $\mathcal{D}$. It splits into two disjoint parts

$$\mathcal{P} = \mathcal{P}_0 \cup \mathcal{P}_1 \,,$$

where $\mathcal{P}_0$ is the set of all 0-dimensional orbits (singular points), and $\mathcal{P}_1$ is the set of all 1-dimensional orbits (singularity-free). We shall here be dealing with the global geometric structure of $\mathcal{P}_1$. To this purpose, first notice that, on the one hand, any $\Gamma \in \mathcal{P}_1$ is a subset of the *carrier* of X

$$\mathrm{carr} X := M - \mathcal{P}_0$$

(which clearly is an open submanifold of M),and being connected in its own manifold topology, Γ is connected in its (coarser) subspace topology too. On the other hand, owing to the Orbital Determinism Theorem, singulatity-free maximal orbits are mutually disjoint and cover $\mathrm{carr} X$. So $\mathcal{P}_1$ is a partition of $\mathrm{carr} X$ into connected subsets. More than that, 1.3.19 **Theorem.** *$\mathcal{P}_1$ is a foliation of carr$X$.*

Proof. $\mathcal{P}_1$ is the phase portrait of singularity-free system $\mathcal{D}_N = (N, X_N)$ with $N := \text{carr}X$. Let V be the 1-dimensional distribution [58] on N spanned by X_N. It is involutive, since for any nowhere-vanishing $Y, Z \in \chi(N)$ belonging to V, we have $Z = fY$ with $f \in C^\infty(N)$ and then the commutator

$$[Y, Z] = [Y, fY] = (Yf)Y + f[Y, Y] = (Yf)Y$$

belongs to V as well. Owing to the consequent integrability of V, we can consider the 1-dimensional foliation $\mathcal{F}$ on N given by the maximal integral manifolds (leaves) of V. We shall prove the theorem by showing that

$$\mathcal{P}_1 = \mathcal{F}.$$

To this end, let us compare – for any $x \in N$ – the unique leaf $\mathcal{L} \in \mathcal{F}$ and the unique orbit $\Gamma \in \mathcal{P}_1$ containing x. Recall that Γ carries a submanifold structure which ensures the smoothness of the mapping $\tilde{\Phi}_x : I_p \to \Gamma$ induced by maximal motion $\Phi_x : I_x \to N$ of $\mathcal{D}_N$. Then, from $\Phi_x = j \circ \tilde{\Phi}_x$ (with $j : \Gamma \hookrightarrow N$) we draw

$$T_x j\big(\dot{\tilde{\Phi}}_x(0)\big) = \dot{\Phi}_x(0) = X_x\,.$$

Since $X_x \neq 0$, $\dot{\tilde{\Phi}}_x(0)$ turns out to be a basis of $T_x\Gamma$ and then, as $X_x \in V_x$,

$$T_x j(T_x\Gamma) = V_x\,.$$

So Γ is an integral manifold of V. As a consequence Γ is an open subset of $\mathcal{L}$ and $\mathcal{L}$ is then union of such open subsets – there is one of them for each point of $\mathcal{L}$ – which either coincide or are disjoint. Since $\mathcal{L}$ is connected, $\mathcal{L} = \Gamma$. ∎

2 Reduction

2.1 Invariant manifolds

Let $\mathcal{D} = (M, X)$ be a dynamical system. A submanifold $\mathcal{L}$ of M such that, for any $x \in \mathcal{L}$, the maximal orbit Γ through x lies in $\mathcal{L}$, is called an *invariant manifold* of $\mathcal{D}$.

2.1.1 **Remark.** The above *invariance condition* on $\mathcal{L}$ does correspond to the invariance of $\mathcal{L}$ under the action of the flow

$$\Phi_t(\mathcal{L} \cap D_t) \subset \mathcal{L} \qquad \forall t \in \mathbb{R}.$$ ∎

Let $\mathcal{L}$ be a smoothness preserving [59], invariant manifold. In this case, for any $x \in \mathcal{L}$, the maximal motion $\Phi_x : I_x \to M$ induces a smooth curve in $\mathcal{L}$, $\tilde{\Phi}_x : I_x \to \mathcal{L}$, characterized by $\Phi_x = j \circ \tilde{\Phi}_x$, where j is the immersion of $\mathcal{L}$ into M. Smoothness allows time derivation of $\tilde{\Phi}_x$, and then

$$\dot{\Phi}_x(0) = T_x j\big(\dot{\tilde{\Phi}}_x(0)\big),$$

that is

$$X_x = T_x j\big(\dot{\tilde{\Phi}}_x(0)\big)$$

with

$$\dot{\tilde{\Phi}}_x(0) \in T_x\mathcal{L}.$$

Hence it follows that

[58] See I 4.2.9.

[59] See I 3.1.

2.1.2 **Lemma.** *If $\mathcal{L}$ is a smoothness preserving, invariant manifold, then X is tangent to $\mathcal{L}$, i.e.,*

$$X_x \in T_x j(T_x \mathcal{L}) \qquad \forall x \in \mathcal{L}.$$ ∎

The above *tangency condition* ammounts to say that there exists, for any $x \in \mathcal{L}$, a unique vector $(X_\mathcal{L})_x \in T_x\mathcal{L}$, such that

$$T_x j(X_\mathcal{L})_x = X_x \, ,$$

i.e., a vector field $X_\mathcal{L} \in \chi(\mathcal{L})$, j-related to X:

$$Tj \circ X_\mathcal{L} = X \circ j.$$

Owing to the canonical isomorphism $T_x\mathcal{L} \simeq T_x j(T_x\mathcal{L})$, for any $x \in \mathcal{L}$, we will put

$$X_\mathcal{L} = X|_\mathcal{L}$$

and will call $X_\mathcal{L}$ the *restriction* of X on $\mathcal{L}$.

2.1.3 **Coordinate expression.** Let $\xi = (x^1, \ldots, x^\ell)$ and $\eta = (x^1, \ldots, x^\ell, c^{\ell+1}, \ldots, c^m)$ be charts on $\mathcal{L}$ and M, distinguished by immersion j. It is easy to check that the components of $X_\mathcal{L}$ in ξ satisfy

$$X^i_\mathcal{L} = X^i \circ j|_\mathcal{U}$$

and their smoothness ensures that of $X_\mathcal{L}$. ∎

Then consider *restricted system* $\mathcal{D}_\mathcal{L} = (\mathcal{L}, X_\mathcal{L})$, j-related to $\mathcal{D}$.

2.1.4 **Theorem.** *The flow of $\mathcal{D}_\mathcal{L}$ is $\{\tilde{\Phi}_x; x \in \mathcal{L}\}$.*

Proof. Let $\{\gamma_x : J_x \to \mathcal{L}; x \in \mathcal{L}\}$ denote the flow of $\mathcal{D}_\mathcal{L}$. On the one hand, for any $x \in \mathcal{L}$, it is

$$Tj \circ \dot{\tilde{\Phi}}_x = \dot{\Phi}_x = X \circ \Phi_x = X \circ j \circ \tilde{\Phi}_x = Tj \circ X_\mathcal{L} \circ \tilde{\Phi}_x$$

and then, Tj being injective at any point of $\mathcal{L}$,

$$\dot{\tilde{\Phi}}_x = X_\mathcal{L} \circ \tilde{\Phi}_x$$

i.e., $\tilde{\Phi}_x$ is a motion of $\mathcal{D}_\mathcal{L}$ (starting from x). Global determinism then ensures that

$$I_x \subset J_x \tag{25}$$

and

$$\tilde{\Phi}_x = \gamma_x|_{I_x} \, . \tag{26}$$

On the other hand, j-relatedness of $\mathcal{D}_\mathcal{L}$ and $\mathcal{D}$ implies that $j \circ \gamma_x : J_x \to M$ is a motion of $\mathcal{D}$ (starting from x), and then a restriction of Φ_x, whence

$$J_x \subset I_x \, . \tag{27}$$

Owing to (25), (27):

$$I_x = J_x \, ,$$

equality (26) finally reads

$$\tilde{\Phi}_x = \gamma_x$$

for all $x \in \mathcal{L}$. ∎

The above theorem shows that, if $\mathcal{L}$ is a smoothness-preserving, invariant manifold, the problem of determining the partial flow $\{\Phi_x; x \in \mathcal{L}\}$ of $\mathcal{D}$ through $\mathcal{L}$, reduces to the problem of determining the flow of $\mathcal{D}_\mathcal{L}$, i.e., to integrating a differential equation on a (generally) lower dimensional manifold.

2.1.5 **Remark.** Notice that $N := \operatorname{carr} X$ is a trivial example of smoothness preserving, invariant manifold of $\mathcal{D}$. So the partial flow $\{\Phi_x; x \in N\}$ of $\mathcal{D}$ through N is (the image under $j : N \hookrightarrow M$ of) the flow of $\mathcal{D}_N$.

2.2 Reductions and constants of the motion

Let

$$\pi : M \to B$$

be a submersion [60] from the phase space M to a smooth manifold B, which *projects* the velocity field X onto a vector field $Y \in \chi(N)$ on the quotient manifold $N := \pi(M) \subset B$, i.e.,

$$T\pi \circ X = Y \circ \pi.$$

In such a case, the submersione π is called a *reduction* of $\mathcal{D} = (M, X)$.
Now consider *reduced system* $\mathcal{D}_\pi := (N, Y)$, π-related to $\mathcal{D}$. If $\mathcal{P}_\pi$ denotes the phase portrait of $\mathcal{D}_\pi$, its lift by π is

$$\pi^*(\mathcal{P}_\pi) := \{\mathcal{L} \subset M \mid \mathcal{L} \text{ is a connected component of } \pi^{-1}(\Gamma_\pi)\,,\ \text{with}\, \Gamma_\pi \in \mathcal{P}_\pi\}.$$

The main result of reduction theory is the following

2.2.1 **Theorem.** *$\pi^*(\mathcal{P}_\pi)$ is a partition of M into smoothness preserving, invariant manifolds.*

Proof. Recall that $\mathcal{P}_\pi$ is a partition of N. Moreover, it is disjoint union of sub-portraits $\mathcal{P}_{\pi_1}$ and $\mathcal{P}_{\pi_0}$ made of maximal orbits which meet or do not meet $\operatorname{carr}Y$, respectively; as is well known, $\mathcal{P}_{\pi_0}$ is a set of singletons and $\mathcal{P}_{\pi_1}$ is a 1-dimensional foliation of the open submanifold $\operatorname{carr}Y \subset \mathcal{N}$ [61]. As a consequence, $\pi^*(\mathcal{P}_\pi)$ is a partition of M. Moreover, it is the disjoint union of $\pi^*\big(\mathcal{P}_{\pi_1}\big)$ and $\pi^*\big(\mathcal{P}_{\pi_0}\big)$. As to $\pi^*(\mathcal{P}_{\pi_0})$, each one of its elements is an open subset of a fibre of π, and then an $(m-n)$-dimensional embedded submanifold of M [62]. As to $\pi^*(\mathcal{P}_{\pi_1})$, it is an $(m-n+1)$-dimensional foliation of the open submanifold $\pi^{-1}(\operatorname{carr}Y) \subset M$, and then each one of its elements is an $(m-n+1)$-dimensional, smoothness preserving submanifold of M [63]. So, what is left, is to prove that each manifold of $\pi^*(\mathcal{P}_\pi)$ is invariant. To this end, let us denote the flows of $\mathcal{D}$ and $\mathcal{D}_\pi$ by

$$\{\Phi_x : I_x \to M; x \in M\} \quad \text{and} \quad \{\Psi_y : J_y \to N; y \in N\}$$

respectively. Now let $\mathcal{L} \in \pi^*(\mathcal{P}_\pi)$ be a connected component of $\pi^{-1}(\Gamma_\pi)$, with $\Gamma_\pi \in \mathcal{P}_\pi$. For any $x \in \mathcal{L}$, π-relatedness of $\mathcal{D}$ and $\mathcal{D}_\pi$ implies that $\pi \circ \Phi_x : I_x \to N$ is a motion of $\mathcal{D}_\pi$, starting from $y := \pi(x) \in \Gamma_\pi = \Psi_y(J_y)$. Then $I_x \subset J_y$ and

$$\pi \circ \Phi_x = \Psi_y|_{I_x}$$

whence

$$\Phi_x(t) \in \pi^{-1}\big(\Psi_y(t)\big) \subset \pi^{-1}(\Gamma_\pi) \qquad \forall t \in I_x$$

---

(60) See I 3.2.
(61) See II 1.3.19.
(62) See Implicit Function Theorem I 3.2.5. Recall that $n = \dim N = \dim B$.
(63) See I 3.2.7.

that is

$$x \in \Gamma := \Phi_x(I_x) \subset \pi^{-1}(\Gamma_\pi)$$

or, owing to the connectedness of Γ

$$\Gamma \subset \mathcal{L}. \qquad \blacksquare$$

As a consequence, in view of the theorem of the previous section, the flow of $\mathcal{D}$ turns out to be the union of the disjoiunt flows of systems

$$\{\mathcal{D}_{\mathcal{L}} = (\mathcal{L}, X_{\mathcal{L}}) \, : \, \mathcal{L} \in \pi^*(\mathcal{P}_\pi)\}.$$

This shows that a reduction π just splits up the problem of 'integrating' $\mathcal{D}$ into two problems (both on lower dimensional manifolds):
(i) first, to integrate reduced system $\mathcal{D}_\pi$ – at least at the level of phase potrait;
(ii) then, to integrate systems $\mathcal{D}_{\mathcal{L}}$'s – which will consequently be called the *residuals* of π.
A fortunate case is one in which the reduced system or the residuals are 'integrable'. A simple example of integrable reduced system is the following.
Let

$$k : M \longrightarrow B$$

be a smooth mapping from the phase space M to a smooth manifold B, which projects the velocity field X onto the null vector field $Y = 0$ of B, i.e.,

$$Tk \circ X = 0.$$

The mapping k is called a *constant of the motion* of $\mathcal{D} = (M, X)$.

2.2.2 **Proposition.** *A smooth mapping $k : M \to B$ is a constant of the motion of $\mathcal{D} = (M, X)$ if, and only if,*

$$k \circ \Phi_x = \text{const.} \qquad \forall x \in M.$$

Proof. Just notice that, for any $x \in M$,

$$\frac{d}{dt}(k \circ \Phi_x) = Tk \circ \dot{\Phi}_x = Tk \circ X \circ \Phi_x \, .$$

So, if $k \circ \Phi_x = \text{const.}$, then $\frac{d}{dt}(k \circ \Phi_x) = 0$ identically and, in particular,

$$(Tk \circ X)_x = \frac{d}{dt}(k \circ \Phi_x)(0) = 0.$$

Conversely, if $Tk \circ X = 0$, then $\dfrac{d}{dt}(k \circ \Phi_x) = 0$ and consequently (owing to the global determinism of first-order equations)

$$k \circ \Phi_x = k(x) = \text{const.} \qquad \blacksquare$$

The case of a constant of the motion k, of constant maximal rank given by

$$\text{rank } k = \dim M$$

is trivial since the existence of such a constant of the motion obviously amounts to say that $X = 0$ and then $\mathcal{D}$ is trivially integrable (all of its maximal motions reducing to rest).

The case of a constant of the motion k, of constant maximal rank given by

$$\text{rank } k = \dim B$$

is, on the contrary, remarkable. Indeed in this case, k is a submersion which projects X onto the null vector field $Y = 0$ on quotient manifold $N := k(M) \subset B$, i.e., a reduction of $\mathcal{D}$ whose reduced system $\mathcal{D}_k = (N, Y)$ – trivially integrable – has a phase portrait

$$\mathcal{P}_k = \mathcal{P}_{k_0} = N.$$

Lift $k^*(\mathcal{P}_k)$ is then the foliation of M determined by submersion k (64). Residuals of k are then the restrictions of $\mathcal{D}$ to the leaves of k – called (connected) *level manifolds* of k. The problem of integrating $\mathcal{D}$ directly reduces to the problem of integrating the residuals on the (m-n)-dimensional level manifolds of the constant of the motion.

2.3 First integrals

Let $f \in C^\infty(M)$ be a real-valued smooth function on phase space M, whose Lie derivative along velocity field X identically vanishes:

$$Xf = 0.$$

Such a function f is called a *first-integral* of $\mathcal{D} = (M, X)$. More generally, consider a system of functions $f^1, \dots, f^n \in C^\infty(M)$ or, equivalently, a smooth mapping

$$F = (f^1, \dots, f^n) : M \longrightarrow \mathbb{R}^n$$

2.3.1 **Proposition.** *Functions $f^1, \dots, f^n \in C^\infty(M)$ are first integrals, if and only if $F = (f^1, \dots, f^n)$ is a constant of the motion.*

Proof. Recall that

$$dF \circ X = (df^1 \circ X, \dots, df^n \circ X) = (Xf^1, \dots, Xf^n)$$

and that, at any $x \in M$, T_xF and d_xF are related to each other by a linear isomorphism (65). So

$$TF \circ X = 0 \iff dF \circ X = 0 \iff (Xf^1, \dots, Xf^n) = 0. \qquad \blacksquare$$

The n-tuple $(f^1, \dots, f^n)$ is said to be an *independent* system of functions, if, for any $x \in M$, $(d_xf^1, \dots, d_xf^n)$ is a linearly independent system of covectors.

2.3.2 **Lemma.** *The family $(f^1, \dots, f^n)$ is an independent system of functions, if and only if $F = (f^1, \dots, f^n)$ is a submersion.*

Proof. Recall that, in an admissible chart ξ at $x \in M$,

$$(\text{rank } F)_x = \text{rank } \left(\frac{\partial f^\alpha}{\partial x^i}\right)_x$$

(64) See I 3.2.7.

(65) See I 2.2.5.

and that the rows of matrix $\left(\dfrac{\partial f^\alpha}{\partial x^i}\right)_x$ are the m-tuples of components in ξ of covectors $(d_x f^\alpha)$. Hence, it follows that F is a submersion at x iff

$$(\text{rank } F)_x = n \iff \text{rank } \left(\frac{\partial f^\alpha}{\partial x^i}\right)_x = n$$
$$\iff (d_x f^1, \ldots, d_x f^n) \text{ is linearly independent.}$$

∎

So, n independent first integrals $(f^1, \ldots, f^n)$ yield a constant of the motion F of constant, maximal rank n, and then they reduce the problem of integrating $\mathcal{D}$ to the integration of the residuals on the $(m-n)$-dimensional level manifolds of F. Then, the more independent first integrals one finds, the lower dimensional residuals will be left to integrate. An extreme example is the following

2.3.3 **Example.** Check that, if $\mathcal{D}$ admits $m = \dim M$ independent first integrals, it is trivially integrable (i.e., all of its maximal motions reduce to rest). ∎

A more meaningful example is given in the following

2.3.4 **Theorem.** *If $\mathcal{D}$ admits m-1 independent first integrals, it is integrable (by a quadrature).*

Proof. Owing to the hypothesis, the problem of integrating $\mathcal{D}$ reduces to the integration of 1-dimensional residuals. Let $\mathcal{D}_\mathcal{L} = (\mathcal{L}, \mathcal{X}_\mathcal{L})$ be a residual. Since $\mathcal{L}$ is a 1-dimensional connected manifold, it is diffeomorphic to $\mathbb{R}$ or S_1 (66). In any case, one can find a covering map

$$h : \mathbb{R} \longrightarrow \mathcal{L}$$

and then a vector field $A \in \chi(\mathbb{R})$, h-related to $\mathcal{X}_\mathcal{L}$,

$$Th \circ A = \mathcal{X}_\mathcal{L} \circ h.$$

As is known, the flow of $\mathcal{D}_\mathcal{L}$ is the h-image of the flow of $(\mathbb{R}, A)$ (67) and this one can be worked out by a quadrature. ∎

2.3.5 **Note.** The global integtration of $(\mathbb{R}, A)$ is an exercise of both calculus and dynamics. We will be concerned with the maximal solution of Cauchy problem $(\mathbb{R}, A, x_0)$, where vector field A will be regarded as a smooth function $A \in C^\infty(\mathbb{R})$, and initial point x_0 will be assumed to be non-singular: $x_0 \in \text{carr}\, A$, say $A(x_0) > 0$. Let W be the connected component (open interval) of $\text{carr}\, A$ containing x_0. As the maximal orbits which meet W are all singularity-free and connected, they lie in W. This means that W is an invariant manifold. So we are led to consider restricted problem (W, A_W, x_0), where A_W only takes positive values. We shall work out its maximal solution by means of the following quadrature:

$$c : x \in W \longmapsto t := \int_{x_0}^{x} \frac{1}{A_W}\, dx \in \mathbb{R}\,.$$

This method defines a smooth function c on W, whose derivative is

$$\frac{dc}{dx} = \frac{1}{A_W} > 0\,,$$

(66) See J.Milnor, *Topology from the differential viewpoint*, (1965).
(67) See II 1.2.5.

which shows that c is an injective local diffeomorphism. As a consequence the image $I := c(W) \subset \mathbb{R}$ is a connected open subset (open interval), containing $c(x_0) = 0$. The induced mapping $\tilde{c} : W \to I$ is a bijective local diffeomorphism, and then a diffeomorphism. Consider inverse diffeomorphism

$$\gamma := \tilde{c}^{-1} : I \longrightarrow W$$

For any $t \in I_{x_0}$,

$$\dot{\gamma}(t) = \left.\frac{d\gamma}{dt}\right|_t = \left(\left.\frac{d\tilde{c}}{dx}\right|_{\gamma(t)}\right)^{-1} = \left(\left.\frac{dc}{dx}\right|_{\gamma(t)}\right)^{-1} = A_W\big(\gamma(t)\big)$$

i.e.,

$$\dot{\gamma} = A_W \circ \gamma$$

and

$$\gamma(0) = x_0 .$$

So γ is a solution of (W, A_W, x_0). Now, maximal solution γ_{x_0} of this problem still defines (owing to its detivative $A_W > 0$) a diffeomorphism – extension of γ – of an open interval onto the same image W as γ's. This implies $\gamma_{x_0} = \gamma$. So γ is the maximal solution of (W, A_W, x_0). ∎

---

3 Hamiltonian Systems

3.1 Symplectic geometry

Let M be an m-dimensional smooth manifold. An exterior 2-form $\omega \in \Lambda_2(M)$ defines a vector bundle *musical* morphism

$$\begin{aligned} {}^\flat : TM &\longrightarrow T^*M \\ X \in T_xM &\longmapsto X^\flat := \omega_x(X, \cdot) \in T_x^*M \end{aligned} \tag{27}$$

and then induces a $C^\infty(M)$-linear mapping [68]

$$\begin{aligned} {}^\flat : \chi M &\longrightarrow \chi^* M \\ X &\longmapsto X^\flat = \omega(X, \cdot) = i_X\omega \end{aligned} \tag{28}$$

If, for each $x \in M$,

$$\operatorname{rank} {}^\flat{}_x = m\,, \tag{29}$$

then the mapping (27) is a vector bundle isomorphism, and the induced mapping (28) is a $C^\infty(M)$-linear isomorphism. In this case, ω is said to be *non-degenerate.*

---

(68) See I 4.4.6.

3.1.1 **Coordinate expression.** Evaluate the coordinate expression of the musical morphism in natural charts ξ^1, ξ_1 on TM and T^*M induced by an admissible chart ξ on M, and check that it is smooth. Also check that the action

$$\alpha = X^\flat$$

(on both vectors and vector fields), is expressed, in terms of components, by

$$\alpha_i = X^j \omega_{ji} = -\omega_{ij} X^j$$

whence one draws that ω is non-degenerate if, and only if, the matrix (ω_{ij}) is non-singular. ∎

A non-degenerate exterior 2-form $\omega \in \Lambda_2(M)$, is said to be an *almost-symplectic structure* on M, and (M, ω) is called an *almost-symplectic manifold.*

3.1.2 **Remark.** If ω is an almost-symplectic structure on M, so is its restriction $\omega|_W$ on any open submanifold $W \subset M$. ∎

3.1.3 **Lemma.** *An almost-symplectic manifold M is even-dimesional.*

Proof. This is an algebraic result. For any $x \in M$, consider the skew-symmetric, bilinear form

$$\omega_x : T_xM \times T_xM \longrightarrow \mathbb{R}.$$

Condition (29) implies that $\omega_x \neq 0$. Then we can find two linearly independent vectors $u_1, v_1 \in T_xM$ such that $\omega_x(u_1, v_1) \neq 0$ or, multiplying one of the vectors by a suitable factor,

$$\omega_x(u_1, v_1) = 1.$$

Let P_1 be the plane spanned by (u_1, v_1) and $E_1 := \{z \in T_xM \,|\, \omega_x(z, u_1) = \omega_x(z, v_1) = 0\}$ the ω_x-orthogonal complement of P_1. Notice that $P_1 \cap E_1 = \{0\}$ (because $\lambda u_1 + \mu v_1 \in E_1 \Rightarrow \lambda = \mu = 0$) and $T_xM = P_1 + E_1$ (because, for any $z \in T_xM$, $z - \big[\omega_x(z, v_1)u_1 - \omega_x(z, u_1)v_1\big] \in E_1$), i.e., $T_xM = P_1 \oplus E_1$. As we can repeat the process on E_1, we choose $(u_2, v_2) \in E_1$ such that

$$\omega_x(u_2, v_2) = 1$$

and we obtain $T_xM = P_1 \oplus P_2 \oplus E_2$, where P_2 is the plane spanned by (u_2, v_2) and E_2 the ω_x-orthogonal complement of P_2 in E_1. If we continue inductively, we finally obtain T_xM as direct sum of a number, say n, of planes

$$T_xM = P_1 \oplus P_2 \oplus \ldots \oplus P_n$$

whence

$$\dim T_xM = 2n.$$ ∎

3.1.4 **Remark.** If we order the above couples $(u_1, v_1), \ldots, (u_n, v_n)$ into one system $(u_1, \ldots, u_n\,;\, v_1, \ldots, v_n)$, we obtain an ω_x-orthonormal basis of T_xM, where ω_x has a matrix of components given by

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

(0 being the null element of $\mathbb{R}^{n^2}$ and 1 the identity of $GL(n, \mathbb{R})$). ∎

On $\mathbb{R}^{2n}$ we have a canonical example of non-degenerate exterior 2-form, globally characterized (in $\xi = \mathrm{id}_{\mathbb{R}^{2n}}$) by the above constant matrix of components.

3.1.5 **Example.** In $\mathbb{R}^{2n}$ let us denote by $(q^h, p_h)_{h=1,\dots,n}$ the natural coordinate functions defined by $\xi = \mathrm{id}_{\mathbb{R}^{2n}}$. Put

$$\Theta := p_h dq^h$$

and

$$\Omega := -d\Theta = dq^h \wedge dp_h \,.$$

From the definition of exterior product we draw

$$\Omega = dq^h \otimes dp_h - dp_h \otimes dq^h = \delta^k_h (dq^h \otimes dp_k) - \delta^h_k (dp_h \otimes dq^k)$$

or

$$\Omega = \Omega_{hk}(dq^h \otimes dq^k) + \Omega^k_h (dq^h \otimes dp_k) + \Omega^h_k (dp_h \otimes dq^k) + \Omega^{hk}(dp_h \otimes dp_k)$$

then

$$(\Omega_{ij}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

whence

$$\det(\Omega_{ij}) = \pm 1$$

according to whether n is even or odd. ∎

Another classical example of a non-degenerate exterior 2-form, locally characterized (in suitable charts) by the above constant matrix of components, is the following

3.1.6 **Example.** Let T^*Q be the cotangent bundle of an n-dimensional smooth manifold Q. Denote by θ_Q the *Liouville* 1-*form* on T^*Q, given, for any $\alpha \in T^*Q$, by

$$\theta_Q(\alpha) := \alpha \circ T_\alpha \pi_Q \;:\; T_\alpha T^*Q \stackrel{T_\alpha \pi_Q}{\longrightarrow} T_{\pi_Q(\alpha)} Q \stackrel{\alpha}{\longrightarrow} \mathbb{R}$$

and let

$$\omega_Q := -d\theta_Q$$

be (up to the sign) its exterior differential . If $\xi = (q^h) : \mathcal{U} \to \mathbb{R}^n$ is an admissible chart on Q, and $\xi_1 = (q^h, p_h) : \mathcal{U}_1 \to \mathbb{R}^n \times \mathbb{R}^n$ the corresponding natural chart on T^*Q, for any $\alpha \in \mathcal{U}_1$, we have

$$\theta_Q(\alpha) = \theta_h(\alpha) d_\alpha q^h + \theta^h(\alpha) d_\alpha p_h$$

with

$$\theta_h(\alpha) := \theta_Q(\alpha)\left(\left.\frac{\partial}{\partial q^h}\right|_\alpha \right) = \alpha\left(T_\alpha \pi_Q \left(\left.\frac{\partial}{\partial q^h}\right|_\alpha \right)\right)$$

and

$$\theta^h(\alpha) := \theta_Q(\alpha)\left(\left.\frac{\partial}{\partial p_h}\right|_\alpha \right) = \alpha\left(T_\alpha \pi_Q \left(\left.\frac{\partial}{\partial p_h}\right|_\alpha \right)\right) .$$

Check that

$$T_\alpha \pi_Q \left(\left.\frac{\partial}{\partial q^h}\right|_\alpha \right) = \left.\frac{\partial}{\partial q^h}\right|_{\pi_Q(\alpha)}$$

$$T_\alpha \pi_Q \left(\left.\frac{\partial}{\partial p_h}\right|_\alpha \right) = 0 \,.$$

As a consequence,

$$\theta_h(\alpha) = \alpha \left. \frac{\partial}{\partial q^h} \right|_{\pi_Q(\alpha)} = \alpha_h = p_h(\alpha)$$
$$\theta^h(\alpha) = 0$$

or,equivalently,

$$\theta_Q(\alpha) = p_h(\alpha) d_\alpha q^h$$

and then

$$\theta_Q|_{\mathcal{U}_1} = p_h dq^h\,.$$

Hence we draw

$$\omega_Q|_{\mathcal{U}_1} = dq^h \wedge dp_h$$

or

$$(\omega_{ij}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

■

Now let (M, ω_M) and (N, ω_N) be almost-symplectic manifolds. A smooth mapping

$$\Psi : M \to N$$

such that

$$\omega_M = \Psi^* \omega_N$$

is said to be a *morphism* (or *isomorphism*, if it is a diffeomorphism) from (M, ω_M) to (N, ω_N). An example of isomorphism is the following.

3.1.7 **Example.** Let X be a vector field on an almost-symplectic manifold (M, ω), such that

$$d_X \omega = 0\,.$$

The above condition is equivalent to the Φ-invariance of ω (where Φ is the flow of X) [69] i.e.

$$\omega|_{D_t} = \Phi_t^* \omega|_{D_{-t}} \qquad \forall t \in \mathbb{R}$$

which means that, for each $t \in \mathbb{R}$, the local transformation $\Phi_t : D_t \to D_{-t}$ is an isomorphism of $(D_t, \omega|_{D_t})$ onto $(D_{-t}, \omega|_{D_{-t}})$.
Such a vector field is called an *infinitesimal automorphism* of (M, ω). ■

(M, ω_M) is said to be *locally isomorphic* to (N, ω_N) if, for any $x \in M$, there exists an isomorphism of an open neighbourhood $(\mathcal{U}, \omega_M|_{\mathcal{U}})$ of x onto an open submanifold $(\mathcal{V}, \omega_N|_{\mathcal{V}})$ of N.

3.1.8 **Example.** Consider (T^*Q, ω_Q). Each $\alpha \in T^*Q$ belongs to the domain of a natural chart ξ_1 which is an isomorphism of $(\mathcal{U}_1\,,\, \omega_Q|_{\mathcal{U}_1})$ onto $\big(\xi(\mathcal{U}) \times \mathbb{R}^n\,,\, \Omega|_{\xi(\mathcal{U}) \times \mathbb{R}^n}\big)$. Then (T^*Q, ω_Q) is locally isomorphic to $(\mathbb{R}^{2n}, \Omega)$. ■

The above example introduces the study of 'integrability'. An almost-symplectic structure ω on a manifold M is said to be *integrable*, if (M, ω) is locally isomorphic to $(\mathbb{R}^{2n}, \Omega)$. This amounts to saying that there exists an atlas of admissible charts on M – the 'local' isomorphisms – where ω is characterized by the same constant matrix of components as Ω's. In this case, (M, ω) is said a *symplectic manifold*, with *symplectic structure* ω and *symplectic charts* which (locally) map ω onto Ω. An integrability condition is given in the following

(69) See II 1.3.11.

3.1.9 **Darboux theorem.** *An almost-symplectic structure ω on a manifold M is integrable if, and only if, it is closed.*

Proof.
(i) If ω is integrable, then for any symplectic chart $(\mathcal{U}, \xi)$, we have

$$\omega|_{\mathcal{U}} = \xi^* \left(\Omega|_{\xi(\mathcal{U})}\right)$$

whence [70]

$$(d\omega)|_{\mathcal{U}} = d\omega|_{\mathcal{U}} = d\xi^* \left(\Omega|_{\xi(\mathcal{U})}\right) = \xi^* d\Omega|_{\xi(\mathcal{U})} = 0$$

(since Ω is exact) and then

$$d\omega = 0.$$

(ii) If ω is closed, for any spherical chart ξ with suitably small domain $\mathcal{U} \subset M$, one can find [71] a diffeomorphism k of the open ball $\xi(\mathcal{U}) \subset \mathbb{R}^{2n}$ onto an open subset $A \subset \mathbb{R}^{2n}$ such that

$$(\xi^{-1})^* \omega|_{\mathcal{U}} = k^* \Omega|_A$$

whence $\omega|_{\mathcal{U}} = \xi^* k^* (\Omega|_A) = (k \circ \xi)^* \Omega|_A$, i.e., $k \circ \xi : \mathcal{U} \to A$ is an isomorphism of $(\mathcal{U}, \omega|_{\mathcal{U}})$ onto $(A, \Omega|_A)$. ∎

3.2 Hamiltonian systems

On an almost-symplectic manifold $(M\omega)$, let us consider the musical isomorphism

$$^\flat : \chi(M) \longrightarrow \chi^*(M)$$

and its inverse

$$^\sharp : \chi^*(M) \longrightarrow \chi(M).$$

They transform exact (and closed) 1-forms onto the following special kinds of vector fields. Let $X \in \chi(M)$ be such that

$$X^\flat = df \qquad (f \in C^\infty(M))$$

or, equivalently,

$$X = df^\sharp \qquad (f \in C^\infty(M)).$$

The vector field X is called a *Hamiltonian field*, and f is said to be a *Hamiltonian function* of X. We will put

$$df^\sharp = X_f\,.$$

More generally, let $X \in \chi(M)$ be such that

$$dX^\flat = 0$$

or, equivalently,

$$X^\flat|_{\mathcal{U}} = df_{\mathcal{U}} \qquad (f_{\mathcal{U}} \in C^\infty(\mathcal{U}))$$
$$X|_{\mathcal{U}} = df_{\mathcal{U}}^\sharp \qquad (f_{\mathcal{U}} \in C^\infty(\mathcal{U}))$$

on each subset $\mathcal{U}$ of an open covering of M. In this case, X is called a *locally Hamiltonian field.*

---

(70) See I 5.3.4.
(71) See R.Abraham and J.E.Marsden, Foundations of Mechanics, (1978), p.175.

3.2.1 **Proposition.** *Any (locally or globally) Hamiltonian field on a symplectic manifold* (M, ω), *is an infinitesimal automorphism of* (M, ω), *and viceversa.*

Proof. Just notice that condition

$$dX^\flat = 0$$

i.e.,

$$di_X\omega = 0,$$

is equivalent, owing to $d\omega = 0$, to

$$(di_X + i_X d)\omega = 0$$

i.e.,

$$d_X\omega = 0.$$

∎

On an almost-symplectic manifold (M, ω), it is possible to give $C^\infty(M)$ an algebra structure defined by skew-symmetric *Poisson brackets*

$$\{f, g\} := \omega(X_f, X_g) \qquad \forall f, g \in C^\infty(M).$$

For any $g \in C^\infty(M)$, Poisson brackets

$$\{\cdot, g\} : f \in C^\infty(M) \mapsto \{f, g\} \in C^\infty(M)$$

act as a Lie derivative, owing to the following

3.2.2 **Proposition.** *For any* $f, g \in C^\infty(M)$,

$$\{f, g\} = X_g f.$$

Proof. Just notice that

$$\{f, g\} = \omega(X_f, X_g) = i_{X_g}(i_{X_f}\omega) = i_{X_g} df = d_{X_g} f.$$

∎

As a consequence the Hamiltonian fields, on a symplectic manifold (M, ω), form a Lie subalgebra of $\chi(M)$.

3.2.3 **Corollary.** *On a symplectic manifold* (M, ω),

$$\begin{cases} dX^\flat = 0 \\ dY^\flat = 0 \end{cases} \Longrightarrow \; [X, Y]^\flat = -d\big(\omega(X, Y)\big)$$

whence

$$[X_f, X_g] = -X_{\{f,g\}}.$$

Proof. Just recall that [72]

$$\begin{aligned} [X, Y]^\flat &= i_{[X,Y]}\omega = (i_X d_Y - d_Y i_X)\omega \\ &= i_X(i_Y d + di_Y)\omega - (i_Y d + di_Y)i_X\omega \\ &= i_X di_Y\omega - i_Y di_X\omega - di_Y i_X\omega \\ &= i_X dY^\flat - i_Y dX^\flat - d\big(\omega(X, Y)\big) \\ &= -d\big(\omega(X, Y)\big) \end{aligned}$$

∎

(72) See I 5.3.7.

3.2.4 **Coordinate expression.** If (M, ω) is a symplectic manifold, in a symplectic chart ξ we have the following specialization of coordinate expression concerning the musical morphisms. Denote the 2n-tuples of components of a vector and a covector (field) by

$$\begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} = \begin{pmatrix} X^h \\ X_h \end{pmatrix} \qquad (h = 1, \ldots, n)$$

$$(\alpha_{(1)}, \alpha_{(2)}) = (\alpha_h, \alpha^h) \qquad (h = 1, \ldots, n)$$

and recall that ω has a matrix of components

$$\begin{pmatrix} \omega_{(11)} & \omega_{(12)} \\ \omega_{(21)} & \omega_{(22)} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

So $\alpha = X^\flat$, or $X = \alpha^\sharp$, iff, in any symplectic chart,

$$\begin{aligned} \alpha_{(1)} &= -\omega_{(11)} X^{(1)} - \omega_{(12)} X^{(2)} = -X^{(2)} \\ \alpha_{(2)} &= -\omega_{(21)} X^{(1)} - \omega_{(22)} X^{(2)} = X^{(1)}, \end{aligned}$$

that is

$$X^{(1)} = \alpha_{(2)} \quad , \quad X^{(2)} = -\alpha_{(1)}$$

or

$$X^h = \alpha^h \quad , \quad X_h = -\alpha_h \, .$$

In particular, $X_f = df^\sharp$ iff

$$X_f^h = \frac{\partial f}{\partial p_h} \quad , \quad X_{f\,h} = -\frac{\partial f}{\partial q^h} \, .$$

As a consequence

$$\{f, g\} = X_g f = \frac{\partial f}{\partial q^h} \frac{\partial g}{\partial p_h} - \frac{\partial f}{\partial p_h} \frac{\partial g}{\partial q^h} \, .$$ ■

Symplectic manifolds and Hamiltonian fields are the ingredients of Hamiltonian dynamics.
A *Hamiltonian system* is a triplet

$$\mathcal{H} = (M, \omega, H)$$

formed by a symplectic manifold (M, ω) and a 'Hamiltonian function' $H \in C^\infty(M)$. Any vector field $X \in \chi(M)$ whose flow leaves ω and H invariant, i.e.,

$$d_X \omega = 0 \quad , \quad d_X H = 0$$

is said to be an infinitesimal automorphism of $\mathcal{H}$. Among the infinitesimal automorphisms of $\mathcal{H}$, there is the Hamiltonian field $X_H = dH^\sharp$, characterized by *Hamilton field equation*

$$i_{X_H} \omega = dH.$$

The dynamical system

$$\mathcal{D}(\mathcal{H}) := (M, X_H)$$

is said to be the dynamical system associated with $\mathcal{H}$. The (maximal) solutions of the *Hamilton equation of motion*

$$\dot{c} = X_H \circ c$$

are called the (maximal) motions of $\mathcal{H}$.

3.2.5 **Coordinate expression.** Let $\xi = (q^h, p_h) : \mathcal{U} \to \mathbb{R}^{2n}$ be a symplectic chart on M. A motion $\gamma : I \to \mathcal{U}$ is a solution of the Hamilton equation of motion, iff

$$\begin{cases} \dfrac{d}{dt}(q^h \circ \gamma) = & \dfrac{\partial H}{\partial p_h} \circ \gamma \\ \dfrac{d}{dt}(p_h \circ \gamma) = & -\dfrac{\partial H}{\partial q^h} \circ \gamma \end{cases}$$

■

A function $f \in C^\infty(M)$ is a first integral of the Hamiltonian system $\mathcal{H} = (M, \omega, H)$, i.e., of the associated dynamical system $\mathcal{D}(\mathcal{H})$, iff [73]

$$X_H f = 0. \tag{30}$$

Owing to the characterization of Poisson brackets as a Lie derivative [74], condition (30) reads

$$\{f, H\} = 0 \tag{31}$$

which is usually expressed by saying that functions f and H are in *involution*. Owing to skew-symmetry of Poisson brackets, condition (31) also reads

$$X_f H = 0 \tag{32}$$

which means that X_f is an infinitesimal automorphism of $\mathcal{H}$. Any Hamiltonian field X_f satisfying condition (32) will be called a *Noether automorphism* of $\mathcal{H}$, and f its *momentum*. The above conditions show that

3.2.6 **Noether Theorem.** *A function $f \in C^\infty(M)$ is a first integral of $\mathcal{H}$, iff it is the momentum of a Noether automorphism of $\mathcal{H}$.* ■

3.3 Hamiltonian reduction

Let $\mathcal{H} = (M, \omega, H)$ be a Hamiltonian system (with $\dim M = 2n$). Let $F := (f^1, \ldots, f^k) : M \to \mathbb{R}^k$ (with $k \leq n$) be a submersion, satisfying

$$\{f^\alpha, f^\beta\} = 0$$

for all $\alpha, \beta = 1, \ldots, k, k+1$ – where $f^{k+1} = H$, i.e., F is an involution set of k independent first integrals of $\mathcal{H}$.

(i) The above involution condition says that any f^β is a first integral of each dynamical system (M, X_{f^α}). This amounts to saying that F is a constant of the motion of (M, X_{f^α}), for all $\alpha = 1, \ldots, k+1$. As a consequence, any leaf S of F is an invariant manifold of each (M, X_{f^α}). This implies that each Hamiltonian field X_{f^α} on M is tangent to S, and then its restriction $X_{f^\alpha} \circ j$ to $S \overset{j}{\hookrightarrow} M$ defines a vector field $X^\alpha \in \chi(S)$ through

$$Tj \circ X^\alpha = X_{f^\alpha} \circ j.$$

(73) See II 2.3.
(74) See II 3.2.2.

As is known, the flow of (M, X_H) through S – $\{\Phi_x; x \in S\}$ – is (the image under j of) the flow of the restricted system $\mathcal{D}_S := (S, X)$, with $X := X^{k+1}$. Remaining systems (S, X^α), with $\alpha = 1, \ldots, k$, will give a contribution – under a suitably hypothesis – to a reduction of $\mathcal{D}_S$.

(ii) We shall now study some geometry on S. First we recall that S (connected component – and then open submanifold – of a fibre of F) is an embedded submanifold of M, whose dimension is

$$\dim S = 2n - k = n + (n - k) \geq n \geq k.$$

For any $x \in S$, the tangent space T_xS will be identified (through T_xj) to vector subspace [75]

$$\ker T_xF = \ker d_xF = \ker(d_xf^1, \ldots, d_xf^k) \subset T_xM.$$

Therefore, we will put

$$X^\alpha = X_{f^\alpha} \circ j.$$

When we consider

$$\omega_S := j^*\omega\,,$$

at any $x \in S$ we shall also put

$$\omega_{S\,x} = \omega_x|_{T_xS \times T_xS}$$

and the musical morphism $^{\flat_S} : TS \to T^*S$ of ω_S will be thought of as acting on each $v \in T_xS$ by

$$v^{\flat_S} = v^\flat|_{T_xS}.$$

3.3.1 **Proposition.** *The characteristic distribution* $\ker{}^{\flat_S}$ *is a k-dimensional, integrable distribution on S.*

Proof. For any $x \in S$, we have

$$\ker{}^{\flat_S}{}_x := \{v \in T_xS \mid v^{\flat_S} = 0\} = T_xS \cap \operatorname{orth}_{\omega_x}(T_xS) \tag{33}$$

where

$$\operatorname{orth}_{\omega_x}(T_xS) := \{v \in T_xM \mid v^\flat|_{T_xS} = 0\}$$

is the ω_x-orthogonal complement of T_xS in T_xM, whose dimension is [76]

$$\dim \operatorname{orth}_{\omega_x}(T_xS) = 2n - (2n - k) = k.$$

Consider now the vector subspace

$$V_x \subset T_xS \tag{34}$$

spanned by $(X^1_x, \ldots, X^k_x)$ in T_xM. As $(X^1_x, \ldots, X^k_x)$ – like $(d_xf^1, \ldots, d_xf^k)$ – is a linearly independent system, we have

$$\dim V_x = k.$$

We easily see that

$$V_x = \operatorname{orth}_{\omega_x}(T_xS). \tag{35}$$

[75] See I 3.2.6.
[76] See C.Godbillon, Géométrie différentielle et Mécanique Analytique, (1969), p.19 (Proposition 1.9).

To this end, owing to the equality of dimensions, it sufficies to check that any

$$v = a_j X^j_x \in V_x$$

belongs to $\mathrm{orth}_{\omega_x}(T_x S)$, i.e.,

$$v^\flat|_{T_x S} = (a_j X^j)^\flat_x|_{T_x S} = a_j ({X^j}^\flat_x)|_{T_x S} = a_j (d_x f^j|_{T_x S}) = 0$$

(for $T_x S = \ker(d_x f^1, \ldots, d_x f^k)$). From (33), (34) and (35) we draw, for any $x \in S$,

$$\ker{}^{\flat_S}{}_x = V_x$$

which proves that $\ker{}^{\flat_S} = V$ is a k-dimensional distribution on S [77]. Now notice that

$$d\omega_S = d(j^*\omega) = j^*(d\omega) = 0$$

and then, for any two vector fields A, B on S belonging to $\ker{}^{\flat_S}$

$$A^{\flat_S} = 0 = B^{\flat_S},$$

we have [78]

$$i_{[A,B]}\omega_S = (i_A d_B - d_B i_A)\omega_S = i_A d_B \omega_S,$$

since $i_A \omega_S = A^{\flat_S} = 0$. As a consequence,

$$i_{[A,B]}\omega_S = i_A(i_B d + d i_B)\omega_S = 0$$

for $i_B \omega_S = B^{\flat_S} = 0$ and $d\omega_S = 0$. In conclusion

$$[A, B]^{\flat_S} = 0$$

which means that

$$[A, B] \in \ker{}^{\flat_S}.$$

This proves that the distribution $\ker{}^{\flat_S}$ is involutive, and then integrable. ∎

As a consequence, $\ker{}^{\flat_S}$ admits maximal integral manifolds which set up a foliation of S, called *characteristic foliation* of ω_S.

(iii) Let us assume that the characteristic foliation of ω_S is a fibration. This amounts to saying that there exists a submersion with connected fibres

$$\rho : S \longrightarrow N$$

such that,

$$\ker T\rho = \ker{}^{\flat_S}.$$

Notice that the fibres of ρ are k-dimensional, and then

$$\dim N = (2n - k) - k = 2(n - k).$$

Moreover the exterior forms

$$\omega_S := j^*\omega \quad , \quad H_S := j^* H$$

turn out to be projectable on N, in view of the following

---

(77) See I 4.2.9.

(78) See I 5.3.7.

3.3.2 **Lemma.** *There exist a unique symplectic structure* ω_N *and a unique Hamiltonian function* H_N *on* N *such that*

$$\begin{aligned} \omega_S &= \rho^*\omega_N \\ H_S &= \rho^* H_N. \end{aligned} \tag{36}$$

Proof. With the aid of admissible charts on S and N, distinguished by ρ, one can see that the existence on N of an almost symplectic structure ω_N and a Hamiltonian function H_N satisfying (36), follows from the invariance of ω_S and H_S under the action of all vector fields belonging to $\ker T\rho$ (i.e., tangent to the fibres of ρ):

$$\begin{aligned} d_Z\omega_S &= 0 \\ d_Z H_S &= 0 \end{aligned}$$

for all $Z \in \chi(S)$ such that $T\rho \circ Z = 0$. The uniqueness of such ω_N and H_N is then ensured by condition (36) itself, and the integrability property $d\omega_N = 0$ follows from $\rho^*(d\omega_N) = d(\rho^*\omega_N) = d\omega_S = 0$. In order to prove the above invariance properties, first notice that, from $\ker T\rho = \ker {}^{\flat_S}$, one draws

$$T\rho \circ Z = 0 \iff i_Z\omega_S = 0$$

and then

$$d_Z\omega_S = i_Z d\omega_S + d i_Z\omega_S = 0.$$

Now notice that, from $\ker T\rho = \ker {}^{\flat_S} = V$, one draws

$$T\rho \circ Z = 0 \iff Z = a_i X^i$$

with $a_j \in C^\infty(S)$, $i = 1, \ldots, k$, and then

$$\begin{aligned} d_Z H_S &= dH^S \circ Z = a_i(dH_S \circ X^i) = a_i(dH \circ Tj \circ X^i) \\ &= a_i(dH \circ X_{f^i} \circ j) = a_i\{H, f^i\} \circ j \\ &= 0 \end{aligned}$$

■

Now we come to the main result

3.3.3 **Theorem.** ρ *is a Hamiltonian reduction, which projects* $\mathcal{D}_S$ *onto the dynamical system associated with* $\mathcal{H}_\rho := (N, \omega_N, H_N)$.

Proof. We have to prove

$$T\rho \circ X = X_{H_N} \circ \rho$$

(recall that $X \in \chi(S)$ is characterized by $Tj \circ X = X_H \circ j$, with $j : S \hookrightarrow M$). To this end, we just start – at each $x \in S$ – from Hamiltonian equation:

$$\begin{aligned}
&\omega_x\big(X_H(x),\cdot\big) = d_x H\\
&\omega_x\big(X_H(x),\cdot\big)\circ T_x j = d_x H \circ T_x j\\
&\omega_x\big(T_x j(X_x), T_x j(\cdot)\big) = d_x(H\circ j)\\
&(j^*\omega)_x(X_x,\cdot) = d_x(j^*H)\\
&\omega_{S\,x}(X_x,\cdot) = d_x H_S\\
&(\rho^*\omega_N)_x(X_x,\cdot) = d_x(\rho^* H_N)\\
&\omega_{N\,\rho(x)}\big(T_x\rho(X_x)\big), T_x\rho(\cdot)\big) = d_x(H_N\circ\rho)\\
&(T_x\rho\cdot X_x)^{\flat_N}\circ T_x\rho = d_{\rho(x)}H_N\circ T_x\rho\\
&(T_x\rho\cdot X_x)^{\flat_N} = d_{\rho(x)}H_N\\
&T_x\rho(X_x) = X_{H_N}\big(\rho(x)\big)\\
&(T\rho\circ X)(x) = (X_{H_N}\circ\rho)(x).
\end{aligned}$$

■

So, by means of k independent first integrals in involution, not only we have a restriction from $\mathcal{H}$ to $\mathcal{D}_S$ (with a dimensional decrease of k units); we also have a Hamiltonian reduction from $\mathcal{D}_S$ to $\mathcal{H}_\rho$ (with a further dimensional decrease of another k units), whose residuals will now be studied.

Let $\mathcal{P}_\rho$ be the phase portrait of $\mathcal{H}_\rho$, and $\rho^*\mathcal{P}_\rho$ its lift by ρ – whose leaves are the phase spaces of the residuals $\mathcal{D}_S = (S, X)$ [(79)]. Let $\mathcal{L}$ be any one of these leaves.

(i) First we shall prove that

3.3.4 **Lemma.** *Each vector field X^α, $(\alpha = 1,\ldots,k,k+1)$, is tangent to $\mathcal{L}$.*

Proof. For $\alpha = k+1$, we have $X^{k+1} = X$ and the result follows from the fact that $\mathcal{L}$ is a smoothness preserving submanifold of S, invariant for $\mathcal{D}_S$. For $\alpha \neq k+1$, notice that, on the one hand, for any $x \in \mathcal{L}$, the fibre $\rho^{-1}(q)$ over $q := \rho(x)$ is a connected subset of $\rho^{-1}(\Gamma)$ – with $q \in \Gamma \in \mathcal{P}_\rho$ – containing x; on the other hand, $\mathcal{L}$ is the connected component of $\rho^{-1}(\Gamma)$ containg x; then $\rho^{-1}(q) \subset \mathcal{L}$. As a consequence, the embedding

$$\iota : \rho^{-1}(q) \hookrightarrow S$$

whose image is contained in $\mathcal{L}$, induces a smooth mapping

$$\tilde{\iota} : \rho^{-1}(q) \hookrightarrow \mathcal{L}$$

such that $\iota = i\circ\tilde{\iota}$, where $i : \mathcal{L} \hookrightarrow S$. Owing to $\ker T_x\rho = V_x$, we have

$$X^\alpha_x \in \ker T_x\rho = T_x\iota\big(T_x\rho^{-1}(q)\big) = T_x i\Big(T_x\tilde{\iota}\big(T_x\rho^1(q)\big)\Big) \subset T_x i(T_x\mathcal{L})$$

which is our claim. ■

So each vector field X^α, restricted to $\mathcal{L}$, yields – through Ti – a vector field Y^α on $\mathcal{L}$

$$Ti\circ Y^\alpha = X^\alpha\circ i.$$

(79) See II 2.2.

As is known, the flow of $\mathcal{D}_S$ through $\mathcal{L}$ is (the image under i of) the flow of residual $\mathcal{D}_{\mathcal{L}} := (\mathcal{L}, Y)$, with $Y := Y^{k+1}$. Remaining systems $(\mathcal{L}, Y^{\alpha})$, with $\alpha = 1, \ldots, k$, will give a contribution – under a suitable hypothesis – to the integration of $\mathcal{D}_{\mathcal{L}}$.

(ii) Now we shall study some geometry on $\mathcal{L}$. Put

$$\ell := \dim \mathcal{L}.$$

If $\mathcal{L} \in \rho^*(\mathcal{P}_{\rho_0})$,

$$\ell = (2n - k) - 2(n - k) = k.$$

If $\mathcal{L} \in \rho^* \mathcal{P}_{\rho_1}$,

$$\ell = (2n - k) - 2(n - k) + 1 = k + 1.$$

In any case

3.3.5 **Lemma.** *The vector fields $(Y^1, \ldots, Y^l)$ are a parallelization of $\mathcal{L}$, i.e., for any $x \in \mathcal{L}$, $(Y^1_x, \ldots, Y^{\ell}_x)$ is a linearly independent system.*

Proof. Let $\ell = k$. For any $x \in \mathcal{L}$, $(Y^1_x, \ldots, Y^k_x)$ is a linearly independent system, because of linear independence of $X^{\alpha}_x = T_x i(Y^{\alpha}_x)$, $\alpha = 1, \ldots, k$. Let $\ell = k + 1$. For any $x \in \mathcal{L}$, put $\rho(x) \in \Gamma \in \mathcal{P}_{\rho}$. On the one hand, from hypothesis $\mathcal{L} \in \rho^*(\mathcal{P}_{\rho_1})$, it follows that

$$\Gamma \in \mathcal{P}_{\rho_1} .$$

On the other hand, from hypothesis per absurdum $Y^{k+1}_x = a_{\alpha} Y^{\alpha}_x$ (with $a_{\alpha} \in \mathbb{R}$ and $\alpha = 1, \ldots, k$), it would follow that

$$X_x = T_x i(Y_x) = a_{\alpha} T_x i(Y^{\alpha}_x) = a_{\alpha} X^{\alpha}_x \in \ker T_x \rho$$

and then

$$(X_{H_N} \circ \rho)(x) = (T\rho \circ X)(x) = T_x \rho(X_x) = 0$$

that is

$$\Gamma \in \mathcal{P}_{\rho_0} .$$

This is in contrast with $\mathcal{P}_{\rho_0} \cap \mathcal{P}_{\rho_1} = \emptyset$. Then $(Y^1_x, \ldots, Y^k_x, Y^{k+1}_x)$ still is a linearly independent system. ∎

3.3.6 **Lemma.** *Vector fields $(Y^1, \ldots, Y^l)$ commute with each other, i.e., for all $\alpha, \beta = 1, \ldots, l$, $[Y^{\alpha}, Y^{\beta}] = 0$.*

Proof. We shall make use of relations $\mathcal{L} \overset{i}{\hookrightarrow} S \overset{j}{\hookrightarrow} M$ relating vector fields $(Y^{\alpha}), (X^{\alpha})$ and $(X_{f^{\alpha}})$, and their Lie brackets too [80]. We have

$$\begin{aligned} T(j \circ i) \circ [Y^{\alpha}, Y^{\beta}] &= Tj \circ Ti \circ [Y^{\alpha}, Y^{\beta}] = Tj \circ [X^{\alpha}, X^{\beta}] \circ i \\ &= [X_{f^{\alpha}}, X_{f^{\beta}}] \circ j \circ i = -X_{\{f^{\alpha}, f^{\beta}\}} \circ (j \circ i) = 0 \end{aligned}$$

As $j \circ i$ is an immersion, we draw $[Y^{\alpha}, Y^{\beta}] = 0$. ∎

The existence of a parallelization on $\mathcal{L}$ set up by commuting vector fields (Y^{α}), is a very peculiar geometrical feature, which strongly recalls the situation of a Euclidean space $\mathbb{R}^{\ell}$ with its natural parallelization defined by commuting vector fields $\left(\frac{\partial}{\partial u^{\alpha}}\right)$ corresponding to natural coordinate function (y^{α}). The only difference is that natural coordinate fields on $\mathbb{R}^{\ell}$ are complete.
From now on we shall assume vector fields $(Y^1, \ldots, Y^l)$ to be complete as well. In this hypothesis, we have

(80) See I 4.2.8 and II 3.2.3.

3.3.7 **Theorem.** *There exists a covering map*

$$h : \mathbb{R}^\ell \longrightarrow \mathcal{L}$$

such that

$$Th \circ \frac{\partial}{\partial y^\alpha} = Y^\alpha \circ h \tag{37}$$

for all $\alpha = 1, \ldots, \ell$.

Proof. We shall first prove that there exists a smooth mapping (local diffeomorphism) satisfying (37). To this end, let Ψ^α be the flow of Y^α (for all $\alpha = 1, \ldots, \ell$). As Y^α is complete, Ψ^α is a one-parameter group of transformations of $\mathcal{L}$. As $[Y^\alpha, Y^\beta] = 0$, Ψ^α and Ψ^β commute with each other [(81)]. As a consequence, if – for any $y = (y^1, \ldots, y^\ell) \in \mathbb{R}^\ell$ and $x \in \mathcal{L}$ – we put

$$\Psi(y, x) = \Psi^1_{y^1} \circ \ldots \circ \Psi^\ell_{y^\ell}(x),$$

we defene an action

$$\Psi : \mathbb{R}^\ell \times \mathcal{L} \longrightarrow \mathcal{L}$$

of additive Lie group $\mathbb{R}^\ell$ on $\mathcal{L}$.
Now, choose a point $x \in \mathcal{L}$ and consider the corresponding trajectory

$$h := \Psi_x : \mathbb{R}^\ell \longrightarrow \mathcal{L}\,.$$

For any $y \in \mathbb{R}^\ell$, the tangent map $T_y h : T_y\mathbb{R}^\ell \to T_{h(y)}\mathcal{L}$ acts on $\left.\frac{\partial}{\partial y^\alpha}\right|_y$ as follows. Notice that $\left.\frac{\partial}{\partial y^\alpha}\right|_y$ is the time derivative at $t = 0$ of

$$\gamma : t \in \mathbb{R} \mapsto \gamma(t) := y + t\delta_\alpha \in \mathbb{R}^\ell$$

(where δ_α is the α-th vector of the canonical basis of $\mathbb{R}^\ell$). Therefore its image

$$T_y h \left(\left.\frac{\partial}{\partial y^\alpha}\right|_y \right)$$

is the time derivative at $t = 0$ of

$$h \circ \gamma : t \in \mathbb{R} \mapsto h\big(\gamma(t)\big) = \Psi_x(y + t\delta_\alpha) \in \mathcal{L}.$$

It is

$$\Psi_x(y + t\delta_\alpha) = \Psi_{y+t\delta_\alpha}(x) = \Psi_y \circ \Psi_{t\delta_\alpha}(x) = \Psi_y \circ \Psi^\alpha_t(x) = \Psi_y \circ \Psi^\alpha_x(t)$$

and then

$$T_y h \left(\left.\frac{\partial}{\partial y^\alpha}\right|_y \right) = (h \circ \gamma)^{\cdot}(0) = T_x\Psi_y \circ \dot{\Psi}^\alpha_x(0) = T_x\Psi_y\big(\dot{\Psi}^\alpha_x(0)\big) = T_x\Psi_y(Y^\alpha_x) \tag{38}$$

Recall that, owing to $[Y^\alpha, Y^\beta] = 0$, Y^α is invariant under each action Ψ^β,

$$T\Psi^\beta_{y^\beta} \circ Y^\alpha = Y^\alpha \circ \Psi^\beta_{y^\beta},$$

(81) See I 1.3.17.

and then it is invariant under action Ψ

$$T\Psi_y \circ Y^\alpha = Y^\alpha \circ \Psi_y\,,$$

whence

$$T_x\Psi_y(Y^\alpha_x) = Y^\alpha_{\Psi_x(y)} = Y^\alpha_{h(y)} \tag{39}$$

From (38) and (39) it follows – for any $y \in \mathbb{R}^\ell$ –

$$T_y h\left(\left.\frac{\partial}{\partial y^\alpha}\right|_y\right) = Y^\alpha_{h(y)}$$

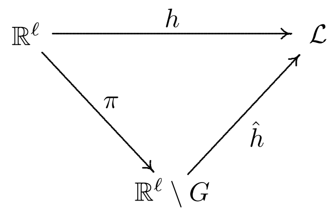
i.e., the statement.
Now we shall only list the main steps that complete the proof [82]. What is left, is to prove that trajectory h is a covering map. To this end, one first check that $h(\mathbb{R}^\ell)$ is a non-void, open and closed subset of connected manifold $\mathcal{L}$, whence

$$h(\mathbb{R}^\ell) = \mathcal{L}.$$

As a consequence, if

$$G := \{y \in \mathbb{R}^\ell \mid h(y) = x\} \subset \mathbb{R}^\ell$$

is the isotropy group of x (Lie subgroup of $\mathbb{R}^\ell$), h induces a bijection $\hat{h}$ of quotient $\mathbb{R}^\ell \setminus G$ onto $\mathcal{L}$ through the commutative diagram



(where π is the canonical projection). The mapping $\hat{h}$ turns out to be a diffeomorphism. Hence we draw that $\dim \mathbb{R}^\ell \setminus G = \ell$ and then $\dim G = 0$, i.e., G is a discrete subgroup of $\mathbb{R}^\ell$. It is known that the only discrete subgroups of $\mathbb{R}^\ell$ are – up to isomorphisms – of the type

$$\underbrace{(0,\ldots,0)}_{r\text{ times}} \times Z^{\ell-r}$$

with $0 \le r \le \ell$. Without loss of generality we put

$$G = \underbrace{(0,\ldots,0)}_{r\text{ times}} \times Z^{\ell-r}$$

and then we have the generalized cylinder

$$\mathbb{R}^\ell \setminus G = (\mathbb{R} \setminus 0)^r \times (\mathbb{R} \setminus Z)^{\ell-r} = \mathbb{R}^r \times T^{\ell-r}$$

[82] For details, see R.Abraham and J.E.Marsden, Foundations of Mechanics, (1978), p.393-394.

where

$$T^{\ell-r} = \underbrace{S_1 \times \ldots \times S_1}_{\ell-r\,\text{times}}$$

is the $(\ell - r)$-dimensional torus. So

$$h : \mathbb{R}^\ell \longrightarrow \mathcal{L}$$

is the composition of the diffeomorphism

$$\hat{h} : \mathbb{R}^r \times T^{\ell-r} \longrightarrow \mathcal{L}$$

with the canonical projection

$$\pi : \mathbb{R}^\ell \longrightarrow \mathbb{R}^r \times T^{\ell-r}$$

which is a classical covering map [(83)]. Therefore h is a covering map. ∎

Finally we come to the problem of integrating $\mathcal{D}_{\mathcal{L}} = (\mathcal{L}, Y)$. Let

$$a_\alpha : \mathcal{L} \longrightarrow \mathbb{R} \qquad (\alpha = 1, \ldots, \ell)$$

be the smooth component of Y with respect to parallelization $(Y^1, \ldots, Y^\ell)$, i.e.,

$$Y = a_\alpha Y^\alpha.$$

3.3.8 **Lemma.** *Y is a parallel vector field, i.e., $a_\alpha = const.$ for all $\alpha = 1, \ldots, \ell$.*

Proof. From Lemma 3.3.2, which holds true also for Y:

$$[Y^\beta, Y] = 0 \qquad (\forall \beta = 1, \ldots, \ell)$$

we draw [(84)]

$$0 = L_{Y^\beta} Y = a_\alpha (L_{Y^\beta} Y^\alpha) + (L_{Y^\beta} a_\alpha) Y^\alpha = (Y^\beta a_\alpha) Y^\alpha$$

or, (Y^α) being linearly independent at each point,

$$0 = Y^\beta a_\alpha = (da_\alpha) Y^\beta \qquad (\forall \alpha, \beta)$$

that is

$$da_\alpha = 0 \qquad\qquad (\forall \alpha)$$

which, $\mathcal{L}$ being connected, proves our claim. ∎

As a consequence, we can consider a simple dynamical system:

$$\mathcal{D} := (\mathbb{R}^\ell, A)$$

with

$$A := a_\alpha \frac{\partial}{\partial y^\alpha}$$

and, making use of covering map h, we see that, for any $y \in \mathbb{R}^\ell$,

$$T_y h(A_y) = a_\alpha T_y h \left(\left. \frac{\partial}{\partial y^\alpha} \right|_y \right) = a_\alpha Y^\alpha_{h(y)} = Y_{h(y)}$$

or equivalently

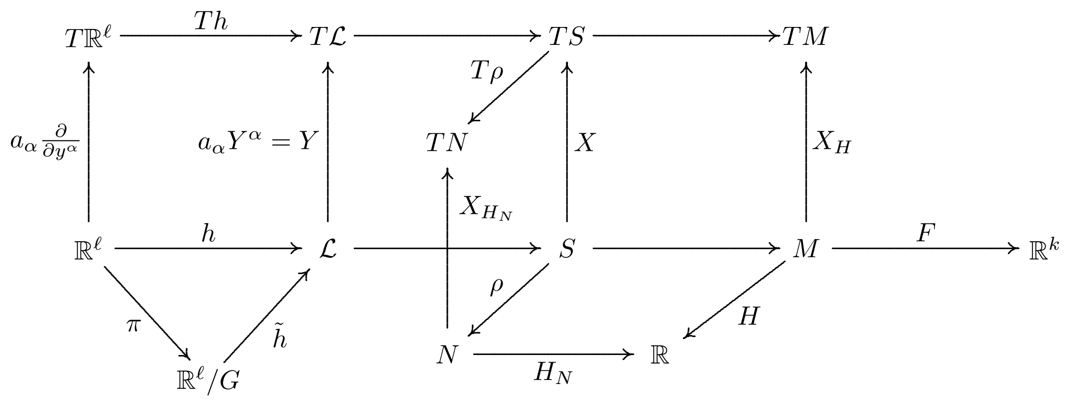
$$Th \circ A = Y \circ h$$

that is

(83) See footnote [(23)].
(84) See II 1.3.14 and 1.3.15.

3.3.9 **Generalized Arnold-Liouville Theorem.**
The dynamical systems $\mathcal{D}$ and $\mathcal{D}_{\mathcal{L}}$ are h-related to each other. ∎

We can summarize the Hamiltonian reduction in the following diagram



The above theorem is the main result of this section. Owing to it, the flow of $\mathcal{D}_{\mathcal{L}}$ comes out to be the image under h of the flow of $\mathcal{D}$, which is trivially known by quadratures. If we put $\mathcal{L} = \mathbb{R}^r \times T^{\ell-r}$, from the *translational* flow of $\mathcal{D}$

$$(t \in \mathbb{R} \mapsto at + y \in \mathbb{R}^\ell \ ; \ y \in \mathbb{R}^\ell)$$

with $a := (a_\alpha)_{\alpha=1,\ldots,l}$, we draw the *helicoidal* flow of $\mathcal{D}_{\mathcal{L}}$:

$$\big(t \in \mathbb{R} \mapsto \pi(at+y) \in \mathbb{R}^r \times T^{\ell-r} \ ; \ \pi(y) \in \mathbb{R}^r \times T^{\ell-r}\big)$$

where

$$\pi(at+y) = \big(a_1 t + y_1, \ldots, a_r t + y_r; a_{r+1} t + y_{r+1} (mod\, 1), \ldots, a_\ell + y_\ell (mod\, 1)\big)$$

The above result then claims that the residuals of reduction ρ of $\mathcal{D}_S$ are integrable by quadratures. This is a generalization of a famous Jacobi-Liouville theorem, which is concerned with the case of $k = n = \frac{1}{2} \dim M$ independent first integrals in involution $F = (f^1, \ldots, f^n)$. In this case, on the one hand, for any level manifold $S \overset{j}{\hookrightarrow} M$ of F, $\dim S = 2n - n = n$; on the other hand $\ker {}^{\flat S}$ is an n-dimensional distribution on S, spanned by the $\mathbb{R}$-linearly independent vector fields (X^α) induced by $(X_{f^\alpha} \circ j)$ $\alpha = 1, \ldots, n$. So $\ker {}^{\flat S} = TS$. Now, TS is a trivially integrable distribution on S, which admits only one maximal integral manifold, actually S itself; therefore, the zero-dimensional space N of all the leaves of characteristic foliation, reduces to a singleton, and then ρ is a trivial reduction from which only one residual arises: $\mathcal{D}_{\mathcal{L}} = kD_S$. As a consequence, if vector fields $(X^1, \ldots, X^n)$ are complete, Theorem 3.3.9 directly shows that $\mathcal{D}_S$ is integrable by quadratures (Jacobi-Liouville theorem).

3.4 Conjugate momenta

Let us turn back to the canonical example of Hamiltonian system $\mathcal{H}$:

$$M = \mathbb{R}^{2n} \quad , \quad \omega = dq^h \wedge dp_h \quad , \quad H = H(q^\beta \,|\, p_\alpha, p_\beta)$$

with ignorable coordinates (q^α), $\alpha = 1, \ldots, k$ [(11)]. It is already known that conjugate momenta $F = (p_\alpha)$ are independent first integrals of $\mathcal{H}$ – whose Hamiltonian fields $d^\sharp p_\alpha)$ are the complete, coordinate vector fields $\left(\frac{\partial}{\partial q^\alpha}\right)$. It is easy to check (e.g., through the coordinate expression of Poisson brackets in a symplectic chart) that conjugate momenta are in involution too. So, if X_H is complete as well, we are in a position like the one assumed in secs 3.1, 3.2. Therefore the problem of integrating $\mathcal{H}$ first reduces to the level manifolds of F – each of which is an affine subspace $S = \{(q^\alpha, q^\beta; \mu_\alpha, p_\beta)\}$ for any choice of constants $\mu_\alpha \in \mathbb{R}$), naturally diffeomorphic to $\mathbb{R}^{2n-k} = \{(q^\alpha, q^\beta; p_\beta)\}$. In S, the characteristic foliation spanned by vector fields $X^\alpha = \left.\frac{\partial}{\partial q^\alpha}\right|_S$ is the fibration of submersion $\rho : S \to N$ (with $N = \mathbb{R}^{2n-2k} = \{(q^\beta; p_\beta)\}$ given by $(q^\alpha, q^\beta; p_\beta) \mapsto (q^\beta; p_\beta)$. So the problem of integrating $(S, X_H|_S)$ furtherly reduces to the integration of Hamiltonian system $\mathcal{H}_\rho$:

$$N = \{(q^\beta; p_\beta)\} \quad , \quad \omega^N = dq^\beta \wedge dp_\beta \quad , \quad H^N(q^\beta; p_\beta) = H(q^\beta; \mu_\alpha, p_\beta)$$

where unknown functions $\big(q^\alpha(t)\big)$ do not appear at all (ignorability). To an orbit $\Gamma \in \mathcal{P}_{\rho 0}$ there corresponds a residual $\mathcal{D}_\mathcal{L}$ whose phase space $\mathcal{L} = \rho^{-1}(\Gamma)$ is a fibre of ρ and then diffeomorphic to $\mathbb{R}^k$, and whose velocity field $X_H|_\mathcal{L}$ has constant components $\left(\left.\frac{\partial H}{\partial p_\alpha}\right|_{\rho^{-1}(\Gamma)}\right)$. To an orbit $\Gamma \in \mathcal{P}_{\rho_1}$ there corresponds a residual $\mathcal{D}_\mathcal{L}$ whose phase space $\mathcal{H} = \rho^{-1}(\Gamma)$ is diffeomorphic to $\mathbb{R}^k \times \mathbb{R}$ or $\mathbb{R}^k \times S_1$, and whose velocity field $X_H|_\mathcal{L}$ can be lifted to the vector field $\frac{\partial}{\partial u^{k+1}}$ on $\mathbb{R}^{k+1}$ by a suitable covering map $h : \mathbb{R}^{k+1} \to \mathcal{L}$.

3.5 Time-dependent systems

A *time-dependent vector field* on a manifold M is a smooth mapping

$$X : \mathbb{R} \times M \longrightarrow TM$$

such that the diagram

$$\begin{array}{ccc} \mathbb{R} \times M & \xrightarrow{\quad X \quad} & TM \\ {\scriptstyle pr_1}\downarrow & & \downarrow{\scriptstyle \tau_M} \\ M & =\!=\!= & M \end{array}$$

is commutative, i.e.,

$$(t, x) \in \mathbb{R} \times M \longmapsto X(t, x) \in T_x M \subset TM\,.$$

X will be identified with the vector field on $\mathbb{R} \times M$ given by

$$\hat{X} : (t, x) \in \mathbb{R} \times M \longmapsto \big(0_t, X(t, x)\big) \in T_t\mathbb{R} \times T_x M \simeq T_{(t,x)}(\mathbb{R} \times M) \subset T(\mathbb{R} \times M)$$

Now, let $\mathbf{t}$ the *time vector field* on $\mathbb{R} \times M$ given by

$$\mathbf{t} : (t, x) \in \mathbb{R} \times M \longmapsto \mathbf{t}(t, x) := \left(\left.\frac{d}{dt}\right|_t, 0_x\right) \in T(\mathbb{R} \times M)$$

(11) See DS III 2.3.

The *suspension* of X is the vector field $\tilde{X}$ on $\mathbb{R} \times M$ given by

$$\tilde{X} := \mathbf{t} + \hat{X},$$

i.e., for any $(t, x) \in \mathbb{R} \times M$,

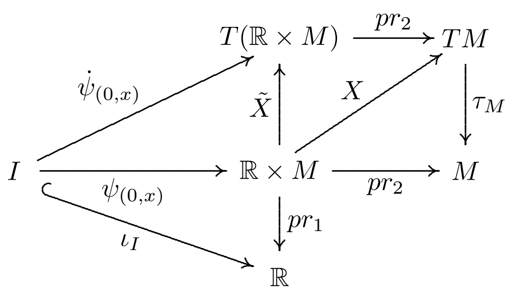
$$\tilde{X}(t, x) = \left(\left.\frac{d}{dt}\right|_t , X(t, x) \right).$$

If $\varphi_x : I_x \subset \mathbb{R} \to M$ is a solution of

$$\begin{cases} \dot{\varphi}_x(t) = X\big(t, \varphi_x(t)\big) \\ \varphi_x(0) = x \end{cases} \tag{23}$$

then φ_x is said to be a *integral curve* of X through x.

3.5.1 **Proposition.** *$\psi_{(0,x)} : I \subset \mathbb{R} \to \mathbb{R} \times M$ is an integral curve of $\tilde{X}$ through $(0, x)$ if, an only is, $pr_1 \circ \psi_{(0,x)} = \iota_I$ and $pr_2 \circ \psi_{(0,x)}$ is an integral curve of X through x*



Proof. Let

$$\psi_{(0,x)}(t) = \big(a(t), b(t)\big)$$

be an integral curve of $\tilde{X}$ through $(0, x)$, i.e., for any $t \in I$,

$$\dot{\psi}_{(0,x)}(t) = \tilde{X}\big(\psi_{(0,x)}(t)\big)$$
$$\big(\dot{a}(t), \dot{b}(t)\big) = \left(\left.\frac{d}{dt}\right|_{a(t)} , X\big(a(t), b(t)\big) \right).$$

As a consequence, from

$$\left.\frac{da}{dt}\right|_t = 1$$

it follows that

$$a(t) = a(0) + t.$$

Owing to the initial condition $a(0) = 0$, it must be $a(t) = t$ and then, for all $t \in I$,

$$\begin{cases} \psi_{(0,x)}(t) = \big(t, b(t)\big) \\ \dot{b}(t) = X\big(t, b(t)\big) \end{cases}$$

i.e.,

$$pr_1 \circ \psi_{(0,x)} = \iota_I$$

and $b(t) = pr_2 \circ \psi_{(0,x)}$ is a solution of (23). Conversely, if $\psi_{(0,x)}(t) = \big(t, \varphi_x(t)\big)$ with $\varphi_x(t)$ solution of (23), then

$$\begin{aligned}\dot{\psi}_{(0,x)}(t) &= \left(\left.\frac{d}{dt}\right|_t, \dot{\varphi}_x(t)\right) = \left(\left.\frac{d}{dt}\right|_t, \big(X \circ \varphi_x\big)(t)\right)\\ &= \tilde{X}\big(t, \varphi_x(t)\big) = \big(\tilde{X} \circ \psi_{(0,x)}\big)(t)\end{aligned}$$

∎

From the above Proposition, one can easily infer an analogous relation between maximal integral curves of $\tilde{X}$ and X.

The *flow* of a complete, time-dipendent vector field X will then be defined by

$$\begin{aligned}\psi :& \mathbb{R} \times M \longrightarrow \mathbb{R} \times M\\ &(t, x) \longmapsto \psi(t, x) := \psi_{(0,x)}(t)\end{aligned}$$

If $X = 0$, i.e., $\tilde{X} = \mathbf{t}$, then $\psi(t, x) = (t, x)$, i.e., $\psi = id_{\mathbb{R}\times M}$.

Let

$$F : \mathbb{R} \times M_1 \longrightarrow \mathbb{R} \times M_2$$

be a fiber bundle isomorphism over $\mathbb{R}$, such that

$$\tilde{X}_1 = F^* \tilde{X}_2 \, .$$

For $\pi \in M_1$, put $F(0, \pi) = (0, x)$. Then, as is well known,

$$F \circ \psi^1_{(0,\pi)} = \psi^2_{(0,x)}.$$

Hence

$$F \circ \psi^1(t, \pi) = \psi^2(t, x).$$

If $\psi^1 = id_{\mathbb{R}\times M_1}$,

$$F(t, \pi) = \psi^2(t, x).$$

3.6 Time-dependent Hamiltonian systems

A *contact manifold* is a pair $(M, \tilde{\omega})$ consisting of a $(2n + 1)$-dimensional manifold M and a closed 2-form $\tilde{\omega}$ of (maximal) rank $2n$ on M. The *characteristic bundle*

$$\mathcal{R}_{\tilde{\omega}} = \{v \in TM : i_v \tilde{\omega} = 0\}$$

is a line bundle (i.e., has 1-dimensional fibres).

3.6.1 **Proposition.** *If (M, ω) is a symplectic manifold, then $(\mathbb{R} \times M, \tilde{\omega})$, with $\tilde{\omega} = pr_2^* \omega$, is a contact manifold. The bundle $\mathcal{R}_{\tilde{\omega}}$ is generated by $\mathbf{t}$, i.e., $i_{\mathbf{t}} \tilde{\omega} = 0$. If $\omega = d\theta$, then $\tilde{\omega} = d\tilde{\theta}$ with $\tilde{\theta} = pr_2^* \theta$.*

Proof. Since ω is closed, $\tilde{\omega}$ is closed too. For any $(t, x) \in \mathbb{R} \times P$,

$$\begin{aligned}\left(r \left.\frac{d}{du}\right|_t, v_x\right) \in \mathcal{R}_{\tilde{\omega}}(t, x) &\iff 0 = \tilde{\omega}_{(t,x)} \left(r \left.\frac{d}{du}\right|_t, v_x\right) = \omega_x(v_x) \circ T_{(t,x)} pr_2\\ &\iff 0 = \omega_x(v_x)\\ &\iff v_x = 0\end{aligned}$$

So

$$\mathcal{R}_{\tilde{\omega}}(t,x) = \left\{ \left(r \left.\frac{d}{du}\right|_t , 0_x \right) : r \in \mathbb{R} \right\} = \operatorname{span} \mathbf{t}(t,x).$$

As a consequence,

$$\operatorname{rank}\ \tilde{\omega}_{(t,x)} = 2n + 1 - \dim \mathcal{R}_{\tilde{\omega}}(t,x) = 2n$$

■

A *time-dependent Hamiltonian system* is a triple (M, ω, H) where (M, ω) is a symplectic manifold and $H \in C^\infty(\mathbb{R} \times M)$. For each $t \in \mathbb{R}$, put

$$\begin{aligned} H_t : M &\longrightarrow \mathbb{R} \\ x &\longmapsto H_t(x) := H(t,x) \end{aligned}$$

(note that $H_t \in C^\infty(M)$) and

$$X_{H_t} := dH_t^\sharp \in \chi(M).$$

Then put

$$\begin{aligned} X_H : \mathbb{R} \times M &\longrightarrow TM \\ (t,x) &\longmapsto X_{H_t}(x) \end{aligned}$$

X_H – or its suspension $\tilde{X}_H \in \chi(\mathbb{R} \times M)$ – is called a *time-dependent Hamiltonian vector field* on (M, ω).

3.6.2 **Remark.** For any $t \in \mathbb{R}$, let

$$\begin{aligned} \lambda_t : M &\longrightarrow \mathbb{R} \times M \\ x &\longmapsto \lambda_t(x) := (t,x) \end{aligned}$$

$$\begin{aligned} T_x\lambda_t : T_xM &\longrightarrow T_t\mathbb{R} \times T_xM \\ v_x &\longmapsto (0_t, v_x) \end{aligned}$$

so that

$$T_x\lambda_t : T_xM \hookrightarrow T_{(t,x)}(\mathbb{R} \times M)$$

Owing to

$$H_t = H \circ \lambda_t\,,$$

we have

$$d_xH_t = d_{(t,x)}H \circ T_x\lambda_t = d_{(t,x)}H|_{\{0_t\} \times T_xM} \tag{3}$$

As a consequence, being $X_H(t,x) = \big(0_t, X_{H_t}(x)\big)$, one has, owing to (3),

$$\begin{aligned} X_HH(t,x) &= \langle d_{(t,x)}H | X_H(t,x)\rangle = \langle d_xH_t | X_{H_t}(x)\rangle \\ &= \omega_x\big(X_{H_t}(x), X_{H_t}(x)\big) = 0. \end{aligned} \tag{4}$$

■

3.6.3 **Remark.**

$$\begin{aligned} i_{X_H}\tilde{\omega}(t,x) &= \omega_x\left(T_{(t,x)}pr_2\big(X_H(t,x)\big), \cdot\right) \circ T_{(t,x)}pr_2 \\ &= \omega_x\big(X_{H_t}(x), \cdot\big) \circ T_{(t,x)}pr_2 \\ &= d_xH_t \circ T_{(t,x)}pr_2 \\ &\overset{(3)}{=} d_{(t,x)}H \circ T_x\lambda_t \circ T_{(t,x)}pr_2 \\ &= d_{(t,x)}H \circ T_{(t,x)}(\lambda_t \circ pr_2). \end{aligned} \tag{5}$$

3.6.4 **Theorem.** *Let* (M, ω, H) *be a time-dependent Hamiltonian system.*
(i) $(\mathbb{R} \times M, \omega_H)$*, with*

$$\omega_H = \tilde{\omega} + dH \wedge dt$$

is a contact manifold. (If $\omega = d\theta$*, then* $\omega_H = d\theta_H$ *with* $\theta_H = \tilde{\theta} + Hdt$*).*
(ii) $\tilde{X}_H$ *is the unique vector field on* $\mathbb{R} \times M$ *satisfying*

$$\begin{cases} i_{\tilde{X}_H}\omega_H = 0 \\ i_{\tilde{X}_H} dt = 1 \end{cases}$$

where $t := pr_1 : \mathbb{R} \times M \to \mathbb{R}$*. (*$\tilde{X}_H$ *generates* $\mathcal{R}_{\omega_H}$*).*
(iii) It is

$$d_{\tilde{X}_H} H = \frac{\partial H}{\partial t}$$

where

$$\frac{\partial H}{\partial t} = d_{\mathbf{t}} H.$$

Proof.
(i)

$$d\omega_H = d\tilde{\omega} + d(dH \wedge dt = d\tilde{\omega} = pr_2^*(d\omega) = 0$$

So ω_H is closed. Let $(t, x) \in \mathbb{R} \times M$ and consider [1]

$$\omega_H(t,x)_{|\{0_t\} \times T_x P|} : \{0_t\} \times T_x P \longrightarrow \big(\{0_t\} \times T_x P\big)^*.$$

(1) Let $\omega : V \longrightarrow V^*$ be a linear mapping and $S \subset V$ a vector subspace. Two more linear mappings can be obtained via restriction:

$$\omega|_S : S \longrightarrow V^* : u \mapsto \omega(u)$$

and

$$\omega_{|S|} : S \longrightarrow S^* : u \mapsto \omega(u)|_S \,.$$

Obviously

$$\ker \omega|_S \subset \ker \omega_{|S|} \,,$$

hence

$$\begin{aligned} \dim\ker \omega|_S &\leq \dim\ker \omega_{|S|} \\ -\dim\ker \omega|_S &\geq -\dim\ker \omega_{|S|} \\ \dim S - \dim\ker \omega|_S &\geq \dim S - \dim\ker \omega_{|S|} \\ \mathrm{rank}\omega|_S &\geq \mathrm{rank}\omega_{|S|} \end{aligned}$$

On the other hand

$$\mathrm{rank}\ \ \omega = \dim \omega(V) \geq \dim \omega(S) = \mathrm{rank}\ \ \omega|_S$$

So

$$\mathrm{rank}\ \ \omega \geq \mathrm{rank}\ \ \omega_{|S|}.$$

We have

$$
\begin{aligned}
\omega_H(t,x)_{|\{0_t\}\times T_xP|}(0_t,v_x) &= \tilde{\omega}(t,x)(0_t,v_x)|_{\{0_t\}\times T_xM} + \big(d_{(t,x)}\mathbf{t}(0_t,v_x)\big)\, d_{(t,x)}H|_{\{0_t\}\times T_xP} \\
&\quad - \big(d_{(t,x)}\mathbf{t}(0_t,v_x)\big)\, d_{(t,x)}H|_{\{0_t\}\times T_xM} \\
&= \tilde{\omega}(t,x)(0_t,v_x)|_{\{0_t\}\times T_xM} \\
&= \omega_x\left(T_{(t,x)}pr_2(0_t,v_x)\right)\circ T_{(t,x)}pr_2|_{\{0_t\}\times T_xM} \\
&= \omega_x(v_x)\circ\iota
\end{aligned}
$$

where $\iota : \{0_t\}\times T_xM \to T_xM$. Hence

$$\text{rank }\ \omega_H(t,x) \geq \text{rank }\ \omega_H(t,x)|_{|\{0_t\}\times T_xM|} = \text{rank }\ \omega(x) = 2n.$$

(ii)

$$
\begin{aligned}
i_{\tilde{X}_H}\omega_H &= i_{\tilde{X}_H}(\tilde{\omega} + dH\wedge dt) \\
&= i_{\tilde{X}_H}\tilde{\omega} + \big(i_{\tilde{X}_H}dH\big)\, dt - \big(i_{\tilde{X}_H}dt\big)\, dH
\end{aligned}
\qquad (*)
$$

Let us consider one term at a time

$$i_{\tilde{X}_H}\tilde{\omega} = i_{(\mathbf{t}+X_H)}\tilde{\omega} = i_{\mathbf{t}}\tilde{\omega} + = i_{X_H}\tilde{\omega}$$

$$i_{\mathbf{t}}\tilde{\omega}(t,x) + \omega_x\big(T_{(t,x)}pr_2\cdot\mathbf{t}(t,x)\big)\circ T_{(t,x)}pr_2 = 0$$

and, owing to (5)

$$i_{X_H}\tilde{\omega} = d_{(t,x)}H\circ T_{(t,x)}(\lambda_t\circ pr_2)$$

so, we have

$$i_{\tilde{X}_H}\tilde{\omega}(t,x) = d_{(t,x)}H\circ T_{(t,x)}(\lambda_t\circ pr_2) \qquad (a)$$

$i_{\tilde{X}_H}dH = i_{(\mathbf{t}+X_H)}dH = i_{\mathbf{t}}dH + i_{X_H}dH = i_{\mathbf{t}}dH = d_{\mathbf{t}}H$ and then

$$\big(i_{\tilde{X}_H}dH\big)\, dt = d_{\mathbf{t}}H dt \qquad (b)$$

$i_{\tilde{X}_H}dt = i_{(\mathbf{t}+X_H)}dt = i_{\mathbf{t}}dt + i_{X_H}dt = 1$ and then

$$\big(i_{\tilde{X}_H}dt\big)\, dH = dH. \qquad (c)$$

As a consequence ([1])

$$i_{\tilde{X}_H}\omega_H(t,x) = d_{(t,x)}H\circ T_{(t,x)}(\lambda_t\circ pr_2) + \langle d_{(t,x)}H|\mathbf{t}(t,x)\rangle dt - d_{(t,x)}dH = 0.$$

The second condition in (ii) is just (7). Since the characteristic bundle is one dimensional, $\tilde{X}_H$ is unique.
(iii) Owing to (6) it is

$$d_{\tilde{X}_H}H = i_{\tilde{X}_H}dH = d_{\mathbf{t}}H.$$

■

([1]) Notice that

$$\langle d_{(t,x)}H|\mathbf{t}(t,x)\rangle dt = d_{(t,x)}H\circ T_{(t,x)}(\lambda_t\circ pr_1)$$

and that

$$T_{(t,x)}(\lambda_t\circ pr_2) + T_{(t,x)}(\lambda_t\circ pr_1) = id_{T_t\mathbb{R}\times T_xM}.$$

3.7 Canonical transformations

Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds. Let

$$\begin{array}{ccc} \mathbb{R} \times M_1 & \xrightarrow{\quad F \quad} & \mathbb{R} \times M_2 \\ {\scriptstyle t_1}\big\downarrow & & \big\downarrow{\scriptstyle t_2} \\ \mathbb{R} & = = = & \mathbb{R} \end{array}$$

with $t_1 = F^* t_2$, be a fibre bundle isomorphism over $\mathbb{R}$. Put, for any $t \in \mathbb{R}$,

$$F_t := pr_2 \circ F \circ \lambda_t : M_1 \longrightarrow M_2 \,.$$

3.7.1 **Theorem.** *The following statements are equivalent*
(i) For any $t \in \mathbb{R}$, F_t is a symplectic diffeomorphism and, for any $H \in C^\infty(\mathbb{R} \times M_2)$, there is a $K \in C^\infty(\mathbb{R} \times M_1)$ s.t.

$$F^* \tilde{X}_H = \tilde{X}_K \,.$$

(ii) There exists a $K_F \in C^\infty(\mathbb{R} \times M_1)$ such that

$$F^* \tilde{\omega}_2 - \tilde{\omega}_1 - dK_F \wedge dt_1 = 0,$$

i.e., if $\tilde{\omega}_1 = -d\tilde{\theta}_1$ and $\tilde{\omega}_2 = -d\tilde{\theta}_2$,

$$d\left(F^* \tilde{\theta}_2 - \tilde{\theta}_1 + K_F dt_1\right) = 0.$$

(iii) There exists a $K_F \in C^\infty(\mathbb{R} \times M_1)$ such that, for all $H \in C^\infty(\mathbb{R} \times M_2)$,

$$F^* \tilde{X}_H = \tilde{X}_K \,,$$

with

$$K := F^* H + K_F \,.$$

F is said to be a canonical transformation if it satisfies one of the above statements.

3.7.2 **Remark.** Let $F = (id_{\mathbb{R}}, f)$, with $f : M_1 \longrightarrow M_2$. Then F is a canonical transformation iff f is a symplectic diffeomorphism.

Proof. From the commutativity of the diagram

$$\begin{array}{ccc} \mathbb{R} \times M_1 & \xrightarrow{\; F = (id_{\mathbb{R}}, f) \;} & \mathbb{R} \times M_2 \\ {\scriptstyle pr_2}\big\downarrow & & \big\downarrow{\scriptstyle pr_2} \\ M_1 & \xrightarrow[\quad f \quad]{} & M_2 \end{array}$$

i.e., $pr_2 \circ F = f \circ pr_2$, it follows that

$$F_t = pr_2 \circ F \circ \lambda_t = f \circ pr_2 \circ \lambda_t = f.$$

So, if F is a canonical transformation, then f is a symplectic diffeomorphism. Conversely, let f be a symplectic diffeomorphism, i.e.,

$$f^*\omega_2 = \omega_1 \,,$$

then, again from the commutativity of the above diagram, we have

$$F^*\tilde{\omega}_2 = F^* pr_2^*\omega_2 = (pr_2 \circ F)^*\omega_2 = (f \circ pr_2)^*\omega_2 = pr_2^* f^*\omega_2 = pr_2^*\omega_1 = \tilde{\omega}_1$$

and then (ii) is satisfied with $K_F = 0$. ∎

Let (M_1, ω_1, H) be a time-dependent Hamiltonian system. A canonical transformation F reduces H to *equilibrium* if, and only if, $K = 0$, i.e.,

$$H \circ F + K_F = 0.$$

In that case, as $X_K = 0$ (i.e. $\tilde{X}_K = \mathbf{t}_1$), one has

$$\Psi_K = id_{\mathbb{R}\times M_1}$$

and then F gives the flow of X_H:

$$\Psi_H(t, x_0) = F(t, \pi_0),$$

with $\pi_0 = F_0^{-1}(x_0)$, $F_0 = pr_2 \circ F \circ \lambda_0$. A method for obtaining such an F is the Hamilton-Jacobi theory, that we are going to discuss in the following, where we will specialize to the case of a nonautonomous Hamiltonian vector field X_H on a cotangent bundle $M_2 := T^*Q_2$.

3.8 Hamilton-Jacobi theory

Let $\dim Q_1 = \dim Q_2$

$$M_1 := T^*Q_1, \qquad \text{with coordinates } (x^h, \pi_h)$$
$$M_2 := T^*Q_2, \qquad \text{with coordinates } (q^h, p_h)$$

Let

$$\begin{aligned} V : \mathbb{R} \times Q_1 \times Q_2 &\longrightarrow \mathbb{R} \\ (t, x, q) &\longmapsto V(t, x, q)\,. \end{aligned}$$

Put

$$\begin{aligned} f : \mathbb{R} \times Q_1 \times Q_2 &\longrightarrow \mathbb{R} \times T^*Q_1 \\ (t, x, q) &\longmapsto (t, x, \pi)\,, \end{aligned}$$

$$\pi := -d_x V_{(t,q)} \in T_x^*Q_1\,,$$

with coordinate expression

$$(t, x^k, q^k) \longmapsto (t, x^h, \pi_h)\,,$$

$$\pi_h := -\frac{\partial V}{\partial x^h}(t, x^k, q^k)\,.$$

Put

$$\begin{aligned} g : \mathbb{R} \times Q_1 \times Q_2 &\longrightarrow \mathbb{R} \times T^*Q_2 \\ (t, x, q) &\longmapsto (t, q, p)\,, \end{aligned}$$

$$p := d_q V_{(t,x)} \in T_q^* Q_2 \,,$$

with coordinate expression

$$(t, x^k, q^k) \longmapsto (t, q^h, p_h) \,,$$
$$p_h := \frac{\partial V}{\partial q^h}(t, x^k, q^k) \,.$$

Assume that f and g be fibre bundle isomorphisms over $\mathbb{R}$ and put

$$F := g \circ f^{-1} : \mathbb{R} \times M_1 \longrightarrow \mathbb{R} \times M_2 \,.$$

3.8.1 **Theorem.** *F is a canonical transformation.*

Proof. We shall show that 5.3.1 (ii) holds. Firstly, we have

$$F^* \tilde{\theta}_2 = (f^{-1})^* \left(g^* (p_h dq^h) \right) = (f^{-1})^* \left(g^* p_h \wedge g^* dq^h \right)$$
$$= (f^{-1})^* \left(p_h \circ g \wedge d(q^h \circ g) \right) = (f^{-1})^* \left(\frac{\partial V}{\partial q^h} \, dq^h \right) .$$

Secondly, we put

$$K_F := (f^{-1})^* \frac{\partial V}{\partial t} \,.$$

Then

$$F^* \tilde{\theta}_2 - \tilde{\theta}_1 + K_F dt = (f^{-1})^* dV = d \left((f^{-1})^* V \right) ,$$

hence our claim. ∎

Now consider a (time-dependent) Hamiltonian system $(T^* Q_2, H)$. Owing to 5.3.1 (iii), one has

$$F^* \tilde{X}_H = \tilde{X}_K \,,$$

with

$$K = F^* H + K_F \,,$$

F being the canonical transformation generated by V, and

$$K_F = (f^{-1})^* \frac{\partial V}{\partial t} \,.$$

F reduces H to equilibrium, i.e., $K = 0$, if, and only if

$$H \circ g \circ f^{-1} + \frac{\partial V}{\partial t} \circ f^{-1} = 0,$$

i.e., V is a solution of the *Hamilton-Jacobi equation* associated with H:

$$H \circ g + \frac{\partial V}{\partial t} = 0,$$

namely,

$$H \left(t, q, d_q V_{t,x} \right) + \frac{\partial V}{\partial t}(t, x, q) = 0,$$

with coordinate expression

$$H \left(t, q^h, \frac{\partial V}{\partial q^h} \right) + \frac{\partial V}{\partial t} = 0.$$

A solution V of the H-J equation, satisfying the above conditions on f and g, is called a *complete integral.* As is known, the flow Ψ of X_H is the transformed by F of the flow $\Phi \circ f$ of X_K. As $X_K = 0$ (i.e. $\tilde{X}_K = \mathbf{t}_1$), the flow of X_K is $id_{\mathbb{R} \times M_1}$. As a consequence, the flow of X_H is given by F itself. Therefore

3.8.2 **Theorem.** *If the H-J equation associated with H admits a complete integral V, the canonical transformation F generated by V gives the flow of (T^*Q_2, H).* ∎

If $H \in C^\infty(T^*Q_2)$, Poincaré's method suggests to look for a complete integral with separate variables (q, t), i.e.,

$$V(t, x, q) = W(x, q) - tE(x).$$

As

$$d_q V_{(t,x)} = d_q\left(W_x - tE(x)\right) = d_q W_x$$

and

$$\frac{\partial V}{\partial t}(t, x, q) = -E(x),$$

the H-J equation reduces to

$$H(q, d_q W_x) = E(x),$$

or

$$H \circ dW_x = E(x),$$

whose coordinate expression is - for all $x \in Q_1$ -

$$H\left(q^h, \frac{\partial W}{\partial q^h}\right) = E\,.$$

3.8.3 **Coordinate expression.**

One has to be aware that in general, generating functions are defined only locally, and indeed, the global theory of generating functions and the associated global Hamilton-Jacobi theory is more sophisticated. Since our goal is to give an introductory presentation of the theory, we will do many of the calculations in coordinates. Recall that in local coordinates, the conditions for a generating function are written as follows. We seek a complete integral of the H-J equation, i.e., a solution of the nonlinear partial differential equation

$$H\left(t, q, \frac{\partial S}{\partial q^h}(t, q, x)\right) + \frac{\partial S}{\partial t}(t, q, x) = 0$$

for the function S relative to the variables (t, q) in an open domain of $\mathbb{R}^{n+1}$ depending parametrically on x in some open domain of $\mathbb{R}^n$. Moreover, we require that

$$\det\left(\frac{\partial^2 S}{\partial q^h \partial x^k}\right) \neq 0$$

in the above-mentioned domains. By the implicit function theorem, this condition is equivalent to local invertibility of the canonical transformation generated by S. Indeed,

$$\begin{cases} p_h = p_h(t, q, x) := \dfrac{\partial S}{\partial q^h}(t, q, x), \\ \pi_h = \pi_h(t, q, x) := -\dfrac{\partial S}{\partial x^h}(t, q, x). \end{cases}$$

The condition on the Jacobian determinant is equivalent to

$$\det\left(\frac{\partial^2 S}{\partial q^h \partial x^k}\right) = \det\left(\frac{\partial p_h}{\partial x^k}\right) \neq 0,$$

which assures that the first equation can be (locally) inverted with respect to x, namely,

$$x^k = x^k(t, q, p).$$

By plugging this equation in the second equation, one gets

$$\pi_k = \hat{\pi}_k(t, q, p) := \pi_h(t, q, x(t, q, p)).$$

The last two equations are the local representatives of the canonical transformation $(t, q, p) \mapsto (t, x, \pi)$. Again, the condition on the Jacobian determinant assures its invertibility (prove it!).

3.8.4 **Remarks.**
(i) In general, the function S develops singularities, or caustics, as time increases, so it must be used with care. This process is, however, fundamental in geometric optics and in quantization. Moreover, one has to be careful with the sense in which S generates the identity at $t = 0$, as it might have singular behavior in t. For example, it is very easy to verify that

$$S(t, q, x) = \frac{1}{2t}|q - x|$$

generates a canonical transformation that is the identity at $t = 0$.
(ii) Here is another link between the Lagrangian and Hamiltonian view of the Hamilton-Jacobi theory. Define S for t close to a fixed time t_0 by the *action integral*

$$S(t, q, x) = \int_{t_0}^{t} L\left(s, q(s), \dot{q}(s)\right) ds,$$

where $q(s)$ is the solution of the Euler-Lagrange equation equaling x at time t_0 and equaling q at time t. One can show that S satisfies the Hamilton-Jacobi equation. See Arnold [1989, Section 4.6] and Abraham and Marsden [1978, Section 5.2] for more information.
(iii) If H is time-independent and W satisfies the time-independent Hamilton-Jacobi equation

$$H\left(q, \frac{\partial W}{\partial q^h}(q, x)\right) = E(x),$$

then $S(t, q, x) = W(q, x) - tE(x)$ satisfies the time-dependent Hamilton-Jacobi equation, as is easily checked. When using this remark, it is important to remember that E is not really a constant, but it equals $H(x, \pi)$, the energy evaluated at (x, π), which will eventually be the initial conditions. We emphasize that one must generate the time t-map using S rather than W. The coordinate expression of the canonical transformation F generated by W reads

$$\begin{cases} q^h = q^h(t, x^k, \pi_k)\,, \\ p_h = \frac{\partial W}{\partial q^h}(x^k, q^k). \end{cases}$$

The first equation is obtained by solving

$$\pi_k = -\frac{\partial W}{\partial x^k}(x^h, q^h) + t\frac{\partial E}{\partial x^k}(x^h)$$

with respect to q^h, while the second equation is evaluated at the same value of q^k of the first one. The flow $\Psi^H_{(q,p)}(t) = F(t, x, \pi)$ has the same expression, where $(x, \pi) := F_0^{-1}(q, h)$ expressed by

$$\begin{cases} x^k = x^k(q^h, p_h), \\ \pi_k = -\dfrac{\partial W}{\partial x^k}(x^h, q^h), \end{cases}$$

where the first equation is obtained by solving

$$\pi_h = \frac{\partial W}{\partial q^h}(x^k, q^k)$$

with respect to x^k, while the second equation is evaluated at the same value of x^h of the first one.
(iv) The Hamilton-Jacobi equation is crucial in the study of the quantum-classical relationship.
(v) The action function S is a key tool used in the proof of the Arnold-Liouville theorem, which gives the existence of action angle coordinates for systems with integrals in involution; see Arnold [1989] and Abraham and Marsden [1978] for details. ■

3.9 The method of separation of variables.

It is sometimes possible to simplify and even solve the Hamilton-Jacobi equation by what is often called the method of separation of variables. Assume that in the Hamilton-Jacobi equation the coordinate q^1 and the term $\partial S/\partial q^1$ appear jointly in some expression $f(q^1, \partial S/\partial q^1)$ that does not involve $q^2, \ldots, q^n, t$. That is, we can write H in the form

$$H(t, q^1, q^2, \ldots, q^n, p_1, p_2, \ldots, p_n) = \hat{H}\left(t, f(q^1, p_1), q^2, \ldots, q^n, p_2, \ldots, p_n\right)$$

for some smooth functions f and $\hat{H}$. Then one seeks a solution of the Hamilton-Jacobi equation in the form

$$S(t, q, x) = S_1(q^1, x^1) + \hat{S}(t, q^2, \ldots, q^n, x^1, \ldots, x^n).$$

We then note that if S_1 solves

$$f\left(q^1, \frac{\partial S_1}{\partial q^1}\right) = C(x^1)$$

for an arbitrary function C and if $\hat{S}$ solves

$$\hat{H}\left(t, C(x^1), q^2, \ldots, q^n, \frac{\partial \hat{S}}{\partial q^2}, \ldots, \frac{\partial \hat{S}}{\partial q^n}\right) + \frac{\partial \hat{S}}{\partial t} = 0,$$

then S solves the original Hamilton-Jacobi equation. In this way, one of the variables is eliminated, and one tries to repeat the procedure. Note that the first equation is an ordinary first order differential equation for S_1 and can be solved by a quadrature. The second equation is again of the H-J form, but with one variable less. When the above procedure can be iterated $n+1$ times, by separating out all space and time variables, the evaluation of a complete integral of the H-J equation reduces to $n+1$ quadratures. In such a case the Hamiltonian system is called *separable.* In fact, a closely related situation occurs when H is independent of time and one seeks a solution of the form

$$S(t, q, x) = W(q, x) + S_1(t).$$

The resulting equation for S_1 has the solution $S_1(t) = -Et$, and the remaining equation for W is the time-independent Hamilton-Jacobi equation. If q^1 is a cyclic variable, that is, if H does not depend explicitly on q^1, then we can choose $f(q^1, p_1) = p_1$, and correspondingly, we can choose $S_1(q^1, x^1) = C(x^1)q^1$. In general, if there are k cyclic coordinates $q^1, q^2, \ldots, q^k$, we seek a solution to the Hamilton-Jacobi equation of the form

$$S(t, q, x) = \sum_{j=1}^{k} C_j(x^j)q^j + \hat{S}\left(q^{k+1}, \ldots, q^n, x^1, \ldots, x^n\right),$$

with $p_i = C_i(x^i)$, $i = 1, \ldots, k$, being the momenta conjugate to the cyclic variables. We note that in order to obtain a separable H-J equation one should choose an appropriate coordinate chart, adapted to the symmetries of the Hamiltonian system under consideration.

3.9.1 **Example: free particle.**

From the Hamiltonian

$$H = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2),$$

one gets

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial q_1}\right)^2 + \left(\frac{\partial S}{\partial q_2}\right)^2 + \left(\frac{\partial S}{\partial q_3}\right)^2\right] + \frac{\partial S}{\partial t} = 0.$$

It is natural to use the method of separation of variables and seek a solution of the form

$$S(t, q_1, q_2, q_3) = X(q_1) + Y(q_2) + Z(q_3) + T(t).$$

The H-J equation reads

$$\frac{1}{2m}\left[\left(\frac{dX}{dq_1}\right)^2 + \left(\frac{dY}{dq_2}\right)^2 + \left(\frac{dZ}{dq_3}\right)^2\right] + \frac{dT}{dt} = 0,$$

whence

$$\frac{dX}{dq_1} = x_1, \qquad \frac{dY}{dq_2} = x_2, \qquad \frac{dZ}{dq_3} = x_3, \qquad \frac{dT}{dt} = -\frac{x_1^2 + x_2^2 + x_3^2}{2m},$$

whose integration yields

$$S(t, q_1, q_2, q_3, x_1, x_2, x_3) = x_1 q_1 + x_2 q_2 + x_3 q_3 - \frac{x_1^2 + x_2^2 + x_3^2}{2m}t.$$

S satisfies the condition of invertibility and generates the canonical transformation

$$x_i = p_i, \qquad \pi_i = -q_i + \frac{x_1}{m}t, \qquad (i = 1, 2, 3).$$

Thus, x_i are the conserved momenta and $-\pi_i$ the initial positions.

3.9.2 Example: harmonic oscillator.

The Hamiltonian of a one dimensional harmonic oscillator is

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

whence the H-J equation reads

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial q}\right)^2 + m^2\omega^2 q^2\right] + \frac{\partial S}{\partial t} = 0.$$

By setting

$$S(t, q, E) = W(q, E) - Et,$$

we get

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial q}\right)^2 + m^2\omega^2 q^2\right] = E,$$

whence

$$W(q, E) = \sqrt{2mE}\int_0^q \sqrt{1 - \frac{m\omega^2 u^2}{2E}}\, du.$$

Now, $W(q, E)$ is the generating function of a canonical transformation from (q, p) to (E, π) where

$$p = \frac{\partial W}{\partial q}, \qquad \pi = -\frac{\partial W}{\partial E}.$$

From the second equation we get

$$\pi = -\frac{1}{2}\sqrt{\frac{2m}{E}}\int_0^q \frac{du}{\sqrt{1 - \frac{m\omega^2 u^2}{2E}}} = -\frac{1}{\omega}\arcsin\left(\sqrt{\frac{m\omega^2}{2E}}\, q\right),$$

whose inverse is

$$q = -\sqrt{\frac{2E}{m\omega^2}}\sin(\omega\pi).$$

The first equation defining the canonical transformation reads

$$p = \sqrt{2mE}\sqrt{1 - \frac{m\omega^2 q^2}{2E}} = \sqrt{2mE}\cos(\omega\pi).$$

We recall that in the new coordinates the Hamiltonian reads

$$K(E, \pi) = H(q(E, \pi), p(E, \pi)) = E$$

and thus E is conserved, while

$$\pi(t) = \pi_0 - t,$$

whence the equations of motion in the original coordinates follow

$$q(t) = \sqrt{\frac{2E}{m\omega^2}}\sin(\omega(t - \pi_0)), \qquad p(t) = \sqrt{2mE}\cos(\omega(t - \pi_0)).$$

We note that (E, π_0) is the image of the initial point (q_0, p_0), namely

$$E = \frac{1}{2m}(p_0^2 + m^2\omega^2 q_0^2) \qquad \pi_0 = -\arctan(m\omega q_0/p_0),$$

thus confirming that E is nothing but the conserved energy of the oscillator.

3.9.3 Example: separable systems in spherical coordinates.

Consider a mass point m in $\mathbb{R}^3$ subjected to conservative forces derived by a potential energy V. The representative of the Hamiltonian function in the cartesian coordinate system reads

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) + V(x, y, z).$$

By introducing spherical coordinates

$$x = r \sin\vartheta \cos\varphi, \qquad y = r \sin\vartheta \sin\varphi, \qquad z = r\cos\vartheta,$$

with $r > 0$, $0 \le \varphi < 2\pi$ and $0 < \vartheta < \pi$, the representative of the Hamiltonian becomes

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\vartheta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2\vartheta}\right) + V(r, \vartheta, \varphi).$$

Suppose that in spherical coordinates the potential has the form

$$V(r, \vartheta, \varphi) = a(r) + \frac{b(\vartheta)}{r^2} + \frac{c(\varphi)}{r^2 \sin^2\vartheta}.$$

Then, the H-J equation

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \vartheta}\right)^2 + \frac{1}{r^2\sin^2\vartheta}\left(\frac{\partial S}{\partial \varphi}\right)^2\right] + V(r, \vartheta, \varphi) + \frac{\partial S}{\partial t} = 0$$

is separable by setting

$$S(t; r, \vartheta, \varphi; \alpha_r, \alpha_\vartheta, \alpha_\varphi) = W_1(\varphi; \alpha_\varphi) + W_2(\vartheta; \alpha_\vartheta, \alpha_\varphi) + W_3(r; \alpha_r, \alpha_\vartheta, \alpha_\varphi) - E(\alpha_r, \alpha_\vartheta, \alpha_\varphi)t.$$

Indeed, by plugging the above Ansatz into the H-J equation one gets

$$\frac{1}{2m}\left(\frac{\partial W_3}{\partial r}\right)^2 + a(r) + \frac{1}{2mr^2}\left\{\left(\frac{\partial W_2}{\partial \vartheta}\right)^2 + 2mb(\vartheta) + \frac{1}{\sin^2\vartheta}\left[\left(\frac{\partial W_1}{\partial \varphi}\right)^2 + 2mc(\varphi)\right]\right\} = E,$$

which splits into a system of first order ordinary differential equations

$$\begin{cases} \left(\dfrac{\partial W_1}{\partial \varphi}\right)^2 + 2mc(\varphi) = e_1(\alpha_\varphi), \\ \left(\dfrac{\partial W_2}{\partial \vartheta}\right)^2 + 2mb(\vartheta) + \dfrac{e_1(\alpha_\varphi)}{\sin^2\vartheta} = e_2(\alpha_\vartheta, \alpha_\varphi), \\ \dfrac{1}{2m}\left(\dfrac{\partial W_3}{\partial r}\right)^2 + a(r) + \dfrac{e_2(\alpha_\vartheta, \alpha_\varphi)}{2mr^2} = E(\alpha_r, \alpha_\vartheta, \alpha_\varphi). \end{cases}$$

The integration of the system yields

$$\begin{cases} W_1 = \displaystyle\int d\varphi \sqrt{e_1(\alpha_\varphi) - 2mc(\varphi)}, \\ W_2 = \displaystyle\int d\vartheta \sqrt{e_2(\alpha_\vartheta, \alpha_\varphi) - 2mb(\vartheta) - \frac{e_1(\alpha_\varphi)}{\sin^2\vartheta}}, \\ W_3 = \displaystyle\int dr \sqrt{2m\left[E(\alpha_r, \alpha_\vartheta, \alpha_\varphi) - a(r) - \frac{e_2(\alpha_\vartheta, \alpha_\varphi)}{2mr^2}\right]}. \end{cases}$$

The above equations particularize in the very important situation of a particle in a central potential $V(r)$. In such a case one can set $c(\varphi) = b(\vartheta) = 0$ and obtain $W_1(\varphi) = p_\varphi \phi$. The coordinate φ is cyclic and $p_\varphi = \pm\sqrt{e_1}$, the z-component of the angular momentum, is conserved. Moreover, $e_2 = p_\varphi^2/\sin^2\vartheta + p_\vartheta^2$ is the square modulus of the angular momentum which is also conserved.

4 Introduction to canonical perturbation theory

4.1 Preliminaries

4.1.1 **Definition.** *A Hamiltonian system is called quasi integrable if the Hamiltonian function is given by*

$$h(q,p,\varepsilon) = h_0(q,p) + \varepsilon h_1(q,p),$$

where $(q,p) \in U$, $0 \leq \varepsilon \leq 1$ *and* h_0 *is the Hamiltonian of a completely integrable system on a time-invariant domain* $U \subset T^*M$. ▪

4.1.2 **Remarks.**
(i) Since h_0 is is the Hamiltonian of a completely integrable system, there exists a canonical transformation to action-angle variables $(q,p) \mapsto (\varphi, J)$, such that the transformed Hamiltonian only depends on J, namely

$$H(\varphi, J, \varepsilon) = H_0(J) + \varepsilon H_1(\varphi, J),$$

with $(\varphi, J) \in T^n \times B$, $B \subset \mathbb{R}^n$ open.
(ii) For $\varepsilon = 0$ the system is integrable and the Hamilton equations read

$$\dot{J} = 0, \qquad \dot{\varphi} = \omega(J),$$

where

$$\omega(J) = \partial_J H_0(J).$$

The phase space T^*M is foliated into invariant tori labeled by actions which are first integrals. The motions are bounded and quasi periodic. Namely

$$J(t) = J(0), \qquad \varphi(t) = \omega(J(0))t + \varphi(0).$$

(iii) When the perturbation is switched on, for $\varepsilon \neq 0$, the action variables are no longer constants of the motion and one gets

$$\dot{J} = -\varepsilon \partial_\varphi H_1(\varphi, J).$$

Therefore,

$$|J(t) - J(0)| < \|\partial_\varphi H_1\| \, \varepsilon t,$$

where $\|f\| = \sup_{(\varphi,J)\in T^n \times B} |f(\varphi, J)|$. ▪

4.1.3 **Remark.** The above is a very crude estimate. It is completely useless for time larger than $O(1)$. In fact it does not take into account the fact that $\partial_\varphi H_1$ is a multiply periodic function of φ with zero mean. ▪

4.1.4 **Example.** The Hamiltonian $H(\varphi, J, \varepsilon) = J + \varepsilon \cos\varphi$, with $(\varphi, J) \in \mathbb{R}^2$ produces the Hamilton equations

$$\dot{J} = \varepsilon \sin\varphi, \qquad \dot{\varphi} = 1,$$

whose solutions are

$$J(t) = J(0) + \varepsilon\left(\cos\varphi(0) - \cos(\varphi(0) + t)\right), \qquad \varphi(t) = \varphi(0) + t.$$

Therefore

$$|J(t) - J(0)| \le 2\varepsilon$$

for all t, and not only for $t = O(1)$. ■

4.2 Perturbation theory

The aim of perturbation theory is to find a canonical transformation that shifts the dependence on φ at order ε^2. Then iterate up to the desired order.

4.2.1 **Remark.** Suppose there exists a canonical transformation that reduces the perturbed Hamiltonian into a completely integrable one. From its generating function $W(\varphi, I, \varepsilon)$ we would get

$$J = \partial_\varphi W(\varphi, I), \qquad \psi = \partial_I W(\varphi, I),$$

that yield the canonical transformation $(\varphi, J) \mapsto (\psi, I)$. Therefore, one would obtain the Hamilton-Jacobi equation

$$H(\varphi, \partial_\varphi W, \varepsilon) = H_0(\partial_\varphi W) + \varepsilon H_1(\varphi, \partial_\varphi W) = \tilde{H}(I, \varepsilon).$$

■

4.2.2 **Remark.** Our requirement is that

$$\tilde{H}(\psi, I, \varepsilon) = \tilde{H}_0(I) + \varepsilon \tilde{H}_1(I) + \varepsilon^2 F(\psi, I, \varepsilon).$$

Therefore we seek for a canonical transformation ε-near to the identity

$$W(\varphi, I, \varepsilon) = I \cdot \varphi + \varepsilon W^{(1)}(\varphi, I)$$

whence

$$H_0(I) + \varepsilon \partial_I H_0(I) \cdot \partial_\varphi W^{(1)}(\varphi, I) + \varepsilon H_1(\varphi, I) + O(\varepsilon^2) = \tilde{H}_0(I) + \varepsilon \tilde{H}_1(I) + O(\varepsilon^2).$$

At zero order in ε

$$\tilde{H}_0(I) = H_0(I).$$

At first order we get

4.2.3 **Definition.** *The fundamental equation of canonical perturbation theory is*

$$\omega(I) \cdot \partial_\varphi W^{(1)}(\varphi, I) + H_1(\varphi, I) = \tilde{H}_1(I),$$

in the unknown functions $W^{(1)}(\varphi, I)$ and $\tilde{H}_1(I)$. ■

4.2.3 **Remarks.**
(i) The fundamental equation is a first order partial differential equation on the torus T^n, and perturbation theory always yields equations of the above form.
(ii) Let us assume that the fundamental equation has a solution $W^{(1)}(\varphi, I)$ and $\tilde{H}_1(I)$. Then, in the new canonical variables we would get from

$$\dot{I} = O(\varepsilon^2), \qquad \dot{\psi} = \omega(I) + \varepsilon \partial_I \tilde{H}_1(I) + O(\varepsilon^2),$$

whence for $t \in [0, \varepsilon^{-1}]$

$$|I(t) - I(0)| = O(\varepsilon).$$

Therefore

$$|J(t) - J(0)| = |J(t) - I(t)| + |I(t) - I(0)| + |I(0) - J(0)| = O(\varepsilon),$$

since $|J(t) - I(t)| = O(\varepsilon)$ uniformly in t. ∎

4.2.4 **Formal solution.**
(i) Since H_1 depends periodically on φ, it can be expanded in a Fourier series:

$$H_1(\varphi, J) = \sum_{k \in \mathbf{Z}^n} \hat{H}_k(J)\, \mathrm{e}^{ik\cdot\varphi},$$

where

$$\hat{H}_k(J) = \int_{T^n} \mathrm{e}^{-ik\cdot\varphi} H_1(\varphi, J)\, \frac{\mathrm{d}^n\varphi}{(2\pi)^n},$$

and analogously for $W^{(1)}$,

$$W^{(1)}(\varphi, J) = \sum_{k \in \mathbf{Z}^n} \hat{W}_k^{(1)}(J)\, \mathrm{e}^{ik\cdot\varphi},$$

with

$$\hat{W}_k^{(1)}(J) = \int_{T^n} \mathrm{e}^{-ik\cdot\varphi} W^{(1)}(\varphi, J)\, \frac{\mathrm{d}^n\varphi}{(2\pi)^n}.$$

(ii) By taking the average of the fundamental equation over T^n, one gets

$$\tilde{H}_1(I) = \hat{H}_0(I),$$

while the other Fourier components, with $k \neq 0$, satisfy the equation

$$ik\cdot\omega(I)\, \hat{W}_k^{(1)}(I) + \hat{H}_k(I) = 0,$$

which is formally solved by

$$W^{(1)}(\varphi, I) = -\sum_{k\neq 0} \frac{\mathrm{e}^{ik\cdot\varphi}}{ik\cdot\omega(I)}\, \hat{H}_k(I). \qquad (*)$$

(iii) Then

$$\tilde{H}(\varphi, I, \varepsilon) = H_0(I) + \varepsilon \hat{H}_0(I) + \varepsilon^2 F(\varphi, I, \varepsilon),$$

where

$$F(\varphi, I, \varepsilon) = \varepsilon^{-2}\left[H_0\left(I + \varepsilon\partial_\varphi W^{(1)}\right) - H_0(I) - \varepsilon\omega\cdot\partial_\varphi W^{(1)}\right] + \varepsilon^{-1}\left[H_1\left(\varphi, I + \varepsilon\partial_\varphi W^{(1)}\right) - H_1(\varphi, I)\right].$$

∎

4.2.5 **Remark.** The denominator of (*) vanishes whenever the frequencies are rationally dependent. But, even when the ratios of the frequencies are irrational, the denominators $k \cdot \omega$ may become arbitrarily small, and the convergence of the formal series (*) must be checked. This is *the problem of small denominators.* ∎

4.2.6 **Theorem.** *Let H_1 be analytic on the domain $|\mathrm{Im}\varphi_j| < \rho$ and $|J_j| < r$, $(j = 1, \ldots, n)$. If there exist $\tau > n-1$ and $c > 0$ such that*

$$|k \cdot \omega| \geq c|k|^{-\tau}, \qquad \forall k \neq 0, \tag{D}$$

where $|k| = |k_1| + \ldots + |k_n|$, $k \in \mathbf{Z}^n$, then the series (*) converges to an analytic function on the same domain , and the following estimates holds:

$$\|W^{(1)}\|_{\rho-\delta,r} \leq \bar{c}\delta^{-\tau-n}\|H_1\|_{\rho,r}, \qquad 0 < \delta < \rho < 2,$$

where $\|f(\varphi,J)\|_{\rho,r} = \sup_{|\mathrm{Im}\varphi_j|<\rho}\sup_{|J_j|<r}|f(\varphi,J)|$ and $\bar{c} = c^{-1}2^{3n}(2\tau/\mathrm{e})^\tau$.

Proof. Since H_1 is periodic in φ, we can shift the path of integration for φ_j so that $\mathrm{Im}\varphi_j = \rho\,\mathrm{sign}k_j$. There results

$$\|\hat{H}_k\|_r = \sup_{|J_i|<r}\left|\int_{T^n} \mathrm{e}^{-ik\cdot\varphi}\mathrm{e}^{-|k|\rho}H_1(\varphi,J)\,\frac{\mathrm{d}^n\varphi}{(2\pi)^n}\right| \leq \mathrm{e}^{-|k|\rho}\|H_1\|_{\rho,r}.$$

This means that

$$\begin{aligned}\|W^{(1)}\|_{\rho-\delta,r} &\leq \sup_{|\mathrm{Im}\,\varphi_i|<\rho-\delta}\sum_{k\neq 0}\left|\mathrm{e}^{ik\cdot\varphi-|k|\rho}\right| c^{-1}|k|^\tau\|H_1\|_{\rho,r}\\ &\leq \sum_{k\neq 0}\mathrm{e}^{-|k|\delta}c^{-1}|k|^\tau\|H_1\|_{\rho,r}.\end{aligned}$$

In order to bound the sum over k, use the inequalities

$$|k|^\tau \leq \left(\frac{2\tau}{\mathrm{e}\delta}\right)^\tau \mathrm{e}^{|k|\delta/2}, \qquad \forall\tau > 0,\ \forall\delta > 0,$$

and

$$\sum_{k\in\mathbf{Z}}\mathrm{e}^{-|k|\delta/2} = \frac{2}{1-\mathrm{e}^{-\delta/2}} - 1 < \frac{8}{\delta}, \qquad 0 < \delta < 2.$$

Together these produce the inequality

$$\begin{aligned}\|W^{(1)}\|_{\rho-\delta,r} &\leq c^{-1}\|H_1\|_{\rho,r}\left(\frac{2\tau}{\mathrm{e}\delta}\right)^\tau\sum_{k\neq 0}\mathrm{e}^{-|k|\delta/2}\\ &\leq c^{-1}\|H_1\|_{\rho,r}8^n\delta^{-\tau-n}\left(\frac{2\tau}{\mathrm{e}}\right)^\tau = \bar{c}\delta^{-\tau-n}\|H_1\|_{\rho,r}.\end{aligned}$$

The same bound for the sum also shows the analyticity. ∎

4.2.7 **Remarks.**
(i) By Cauchy's theorem, the above estimate implies

$$\left\| \frac{\partial^2 W^{(1)}}{\partial I_i \partial \varphi_j} \right\|_{\rho-\delta, r-\gamma} \leq \bar{c}\delta^{-\tau-n+1}\gamma^{-1}\|H_1\|_{\rho,r}.$$

Therefore the matrix

$$\left(\frac{\partial \psi_i}{\partial \varphi_j}\right) = (\delta_{ij}) + \varepsilon \left(\frac{\partial^2 W^{(1)}}{\partial I_i \partial \varphi_j}\right)$$

is invertible for small enough ε, as it is required for the generating function of a canonical transformation.
(ii) The condition on the ω's means that they must be sufficiently rationally independent. When they have rational ratios, resonance behavior occurs, which can amplify the effect of the perturbation dramatically. However, this can happen also in the vicinity of a resonance.
(iii) It is possible to show that for $\tau > n-1$ the measure of the set of frequencies that do not satisfy the Diophantine condition (D),

$$M_{c,\tau} = \left\{\omega \in \mathbb{R}^n : \exists k \in \mathbf{Z}^n \backslash \{0\} \text{ such that } k \cdot \omega < c|k|^{-\tau}\right\}$$

approaches zero as c in any bounded set, that is

$$M_{c,\tau} \cap \{\omega \in \mathbb{R}^n : |\omega| < K\} = O(c), \qquad \forall K > 0,$$

although it contains all rational points, $\mathbf{Z}^n \subset M_{c,\tau}$. Thus $M_{c,\tau}$ is a strange example of an open, dense set of small measure. Its complement, the set of Diophantine frequencies $\mathbb{R}^n \backslash M_{c,\tau}$ is an example of a Cantor set: a nowhere dense, closed set with no interior points, of large measure.

4.2.8 **Outlook.** A perturbed integrable system has been transformed into another integrable system up to order ε^2. The question arises of whether this procedure can be repeated to eliminate the perturbation completely. In fact, there is a fundamental obstruction to this procedure due to a theorem by Poincaré: For a generic analytic Hamiltonian with an arbitrary small perturbation all constants other than H are destroyed. Therefore, there has long been a wide-spread opinion that it is sufficient a "speck of dust" for making the trajectory winding around densely through the energy surface (ergodic system). Thanks to the work of Kolmogorov, Arnold, and Moser, the famous *KAM theory*, it is now known that it is not so. If an integrable system is perturbed, many of the invariant tori are completely destroyed, while others are only deformed. However, if the perturbation is sufficiently small, the ones that are only deformed (named Cantori) fill up most of the phase space. Therefore, even if there exist no constants other than H, for small ε, enough n-dimensional submanifolds exist so that in most cases the system acts virtually like an integrable system.

References

R. Abraham and J.E.Marsden, *Foundations of Mechanics*, Benjamin, Mass., 1978.
F. Brickell and R.S.Clark, Differentiable Manifolds.
B. O'Neill, Semi-Riemannian Geometry
M. Boothby *An introduction to differentiable manifolds and Riemmanian geometry.*
A. Frolicher and A.Nijenhuis *Theory of vector-valued differential forms*, Ind. Math., 18 (1956), pp. 338-385.
Y. Choquet-Bruhat and C.De Witt-Morette *Analysis, Manifolds and Physics*, Part.I, (1982). H.B. Lawson 'Foliations' Bull. Am. Math. Soc., vol.80, n.3 (1974), p.370 and 373.
S. Lang *Differential Manifolds*, (1972).
V.I. Arnol'd, V.V. Kozlov, and A.I. Neishtadt, *Dynamical Systems III*, Encyclopedia of Mathematical Sciences, Springer, New York, 1988.