Lie Groups and Symmetry

Francesco Fassò
Università di Padova
Dipartimento di Matematica "Tullio Levi-Civita"
fasso@math.unipd.it

December 2022

Indice

(Version 2022.1)

Chapter 1

Lie Groups

1.1 Definition and examples

1.1.A Definition. In the following, differentiable or smooth means C^∞ and submanifold means embedded submanifold. Unless otherwise stated all objects are smooth. We will denote the identity element of a group G by e_G or simply by e; for groups of matrices we will usually denote it $\mathbb{I}$ (the unit matrix).

Definition 1.1.1 *A Lie group G is a group which has also the structure of a (real) smooth manifold such that the group product*

$$\mu : G \times G \to G\,, \quad (g,h) \mapsto gh$$

and the group inversion

$$i : G \to G\,, \quad g \mapsto g^{-1}$$

are smooth maps. The dimension of a Lie group G is its dimension as a smooth manifold.

> **Remark:** A topological space may have different, non-diffeomorphic, manifold structures compatible with the given topology (a well known example is $\mathbb{R}^4$ with the Euclidean topology, which has such 'exotic' manifold structures). It so happens that for a group endowed with a (locally Euclidean) topology with respect to which product and inverse are continuous, instead, there is a *unique* smooth structure compatible with that topology. In other words, topological groups are automatically (and in a unique way) Lie groups.

Exercises 1.1.1 *(i) In some books, the definition of Lie groups requires only the smoothness of the map $G \times G \to G$, $(g,h) \mapsto gh^{-1}$. Show that this implies smoothness of product and inverse.*

1.1.B Examples. We give now a number of classical examples of Lie groups. In most cases, we will have to show that a subgroup of a given Lie group is also a submanifold of it—and hence the product and inversion are automatically smooth being the restriction of smooth operations to a submanifold. In order to do that, we will use the following fact: *Let $f : M \to N$ be a smooth map between two manifolds M and N; if, for a given $n \in f(M)$, f has constant rank k in all points of $P = f^{-1}(n)$, then P is a submanifold of M of codimension k.*[1] In this situation, the restriction to P of any smooth map on M is smooth with respect to the submanifold structure of P.

1. $(\mathbb{R}^n, +)$, $n \geq 1$, with the (additive) group structure and the differentiable structure given by it being a vector space. Indeed, the sum $(x, y) \mapsto x + y$ and the group inverse $x \mapsto -x$ are differentiable maps. The dimension is n.

2. The multiplicative group $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$, since both maps $(x, y) \mapsto xy$ and $x \mapsto 1/x$ are differentiable. Its dimension is 1.

3. $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$ is a group with product given by the multiplication of complex numbers. It becomes a Lie group of dimension 2 if $\mathbb{C}^*$ is equipped with the structure of *real* two-dimensional manifold provided by its identification with $\mathbb{R}^2 \setminus 0$ (through the map $z \mapsto (\Re(z), \Im(z))$).

4. The unit circle

$$S^1 \;=\; \big\{z \in \mathbb{C} \,:\, |z| = 1\big\} \sim \big\{x \in \mathbb{R}^2 \,:\, \|x\| = 1\big\}$$

is a subgroup of $\mathbb{C}^*$. It is also a 1-dimensional submanifold of $\mathbb{C}^*$. Hence, it is a Lie group of dimension 1.

5. Direct product of Lie groups. Consider two Lie groups G_1 e G_2. The Cartesian product $G_1 \times G_2$ is a group with product $(g_1, g_2)(g_1', g_2') = (g_1 g_1', g_2 g_2')$. It is also a manifold of dimension $\dim G_1 + \dim G_2$ with the product manifold structure. Smoothness of product and inverse follows from the fact that, in the product manifold structure, a map is smooth whenever its components are smooth.

6. The 1-dimensional torus $\mathbb{T}^1 := \mathbb{R}/2\pi\mathbb{Z}$ and the n-dimensional torus $\mathbb{T}^n := S^1 \times \ldots \times S^1$ (n factors). As all other examples so far, these are abelian Lie groups.

[1] The rank of f at a point m is the rank of the linear map $T_m f : T_m M \to T_{f(m)} N$. If $\operatorname{rank} T_m f = \dim N$ then $T_m f$ is surjective and f is said to be *submersive* or a *submersion* at m. Those points at which a map is submersive are called *regular points*; the *regular level sets* of a map consist entirely of regular points.

7. In the following, we will denote by $\mathrm{L}(n,\mathbb{R})$, or simply $\mathrm{L}(n)$, the set of all $n\times n$ matrices with real entries. $\mathrm{L}(n)$ is a real vector space of dimension n^2: a basis is formed by the n^2 matrices e_{ij} which have all entries zero except a 1 in position (i,j). A matrix $A\in\mathrm{L}(n)$ with entries A_{ij} can then be written as

$$A \;=\; \sum_{ij} A_{ij}e_{ij} \tag{1.1.1}$$

and, as in any vector space, all questions about continuity or smoothness of maps reduce to control that they are continuous or smooth functions of the matrix entries.

The subset of $L(n,\mathbb{R})$ of all invertible matrices, $\mathrm{GL}(n,\mathbb{R})\equiv\mathrm{GL}(n)=\{A\in L(n,\mathbb{R})\;:\;\det A\neq 0\}$, is a (non-abelian, if $n>1$) group with the matrix multiplication as product. Since the function $\det:\mathrm{L}(n)\to\mathbb{R}$ is continuous (it is a polynomial in the matrix entries), $\mathrm{GL}(n)=\det^{-1}(\mathbb{R}\setminus\{0\})$ is an open subset of $L(n)$ and hence an n^2-submanifold of $\mathrm{L}(n)$. The matrix multiplication is polynomial in the matrix entries and hence smooth. The matrix inversion $A\mapsto A^{-1}$ is smooth as well because, by the cofactor expansion (or Cramer's rule), the entries A^{-1} can be written as ratios of two polynomials in the entries of A, with denominator $\det A\neq 0$. Thus, $\mathrm{GL}(n,\mathbb{R})$ is a Lie group of dimension n^2.

8. Similar arguments show that $\mathrm{GL}(n,\mathbb{C})$ is a Lie group of dimension $2n^2$ (it is an open subset of $\mathrm{L}(n,\mathbb{C})$, which is a vector space of real dimension $2n^2$).

9. The orthogonal group $\mathrm{O}(n)\;:=\;\big\{R\in\mathrm{L}(n,\mathbb{R})\;:\;RR^T=\mathbb{I}\big\}$ is a Lie group of dimension $\frac{1}{2}n(n-1)$. Indeed, on the one hand, $\mathrm{O}(n)$ is a subgroup of $\mathrm{GL}(n)$. On the other hand, as we show below, $\mathrm{O}(n)$ is a submanifold of $\mathrm{L}(n)$ of dimension $\frac{1}{2}n(n-1)$; therefore, smoothness of the group operations follows from that of $\mathrm{GL}(n)$.

In order to prove that $\mathrm{O}(n)$ is a submanifold of $L(n)$ we proceed as follows.

- Let $\mathrm{Symm}(n)$ be the vector subspace of $\mathrm{L}(n)$ formed of all symmetric matrices. Since a symmetric matrix has at most $1+2+\ldots+(n-1)+n=\frac{1}{2}n(n+1)$ distinct entries, $\mathrm{Symm}(n)$ has dimension $\frac{1}{2}n(n+1)$.
- Consider the map

$$f:\mathrm{L}(n)\to\mathrm{Symm}(n)\,,\quad A\mapsto AA^T\,, \tag{1.1.2}$$

which is well defined given that $(AA^T)^T=AA^T$ and is smooth because, again, it is polynomial in the matrix entries. Since $A\in\mathrm{O}(n)$ if and only if $f(A)=\mathbb{I}$,

$$\mathrm{O}(n)\;=\;f^{-1}(\mathbb{I})\,.$$

In order to prove that $\mathrm{O}(n)$ is a submanifold of $\mathrm{L}(n)$, we thus prove that f is a submersion at all points of $\mathrm{O}(n)$, namely, that if $R\in\mathrm{O}(n)$ then the tangent map $T_R f : T_R \mathrm{L}(n) \to T_{f(R)}\mathrm{Symm}(n)$ is surjective. This will also prove that $\dim \mathrm{O}(n) = \dim \mathrm{L}(n) - \dim \mathrm{Symm}(n) = n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1)$.

- We begin by computing the tangent map

$$T_A f \ : \ T_A \mathrm{L}(n) \ \to \ T_{f(A)}\mathrm{Symm}(n)$$

at the points $A \in \mathrm{L}(n)$. Instead of determining the entries of this map, which requires choosing coordinates and computing partial derivatives, it is often easier to directly determine the way it transforms tangent vectors—and this is indeed the procedure we will use most of the times. In order to explain how to do it, we first remind the definition of tangent map. The tangent vectors to a manifold M in a point $m \in M$ are the derivatives of curves through m: $v \in T_m M$ if and only if $v = \gamma'(0)$ for a curve $\gamma : I \subset \mathbb{R} \to M$ such that $\gamma(0) = m$. If $f : M \to N$ is a smooth map between two manifolds M and N, then $T_m f : T_m M \to T_{f(m)}N$ is the map that transforms $v = \gamma'(0)$ into $T_m f \cdot v := (f \circ \gamma)'(0)$ (note that this is a linear map between the vector spaces $T_m M$ and $T_{f(m)}N$). If, as in the present case, the manifold M is a vector space E then its tangent spaces can be identified with E and a convenient choice of a curve through a point $x \in E$ and tangent to a vector $v \in E$ is the straight line $t \mapsto x + tv$.

- Proceeding in this way, we consider two matrices $A, V \in \mathrm{L}(n)$ and compute

$$\begin{aligned} T_A f \cdot V &= \frac{d}{dt}\big[f(A+tV)\big]_{t=0} \\ &= \frac{d}{dt}\big[(A+tV)(A+tV)^T\big]_{t=0} \\ &= \frac{d}{dt}\big[AA^T + t(VA^T + AV^T) + t^2 VV^T\big]_{t=0} \\ &= VA^T + AV^T \,. \end{aligned}$$

Since $T_{f(A)}\mathrm{Symm}(n)$ can be identified with $\mathrm{Symm}(n)$, this shows that surjectivity of $T_A f$ amounts to the fact that, for any symmetric matrix S, there exists a matrix $V \in \mathrm{L}(n)$ such that

$$VA^T + AV^T = S \,.$$

This does in fact happen if $A \in \mathrm{O}(n)$ (take $V = \frac{1}{2}SA$). We conclude that $f : \mathrm{L}(n) \to \mathrm{Symm}(n)$ is a submersion at all points of $\mathrm{O}(n)$ and hence $\mathrm{O}(n)$ is a submanifold of $\mathrm{L}(n)$.

We have so far proven that $\mathrm{O}(n)$ is a Lie group of dimension $\frac{1}{2}n(n-1)$. It is called the *orthogonal group*. Let us see some properties of it.

Orthogonal matrices have determinant ± 1 (given that $RR^T = \mathbb{I}$). $\mathrm{O}(n)$ is thus not connected, being the union of the two (closed and nonempty) disjoint subsets $\mathrm{SO}_{\pm}(n) = \{R \in \mathrm{O}(n) \ : \ \det R = \pm 1\}$. It can be proven that $\mathrm{SO}_{\pm}(n)$ are both connected; we will do this later, but only in the case $n = 3$.

Proposition 1.1.2 $\mathrm{O}(n)$ *is compact.*

Proof. Since $\mathrm{L}(n)$ is a finite-dimensional real vector space, by the Heine-Borel theorem a subset of $\mathrm{L}(n)$ is compact if and only if it is closed and bounded (in some norm—they are all equivalent). $\mathrm{O}(n)$ is closed because the function f as in (1.1.2) is continuous. Consider the norm $\|\cdot\|_F$ on $\mathrm{L}(n)$ associated to the (Frobenius) inner product $(A,B)\mapsto \mathrm{tr}\,(A^TB)$. If $A_1,\dots,A_n\in\mathbb{R}^n$ are the columns of A, then $\|A\|_F^2=\sum_i\|A_i\|^2$, where $\|\cdot\|$ is the Euclidean norm. If $A\in O(n)$, its columns are orthonormal vectors and $\|A\|_F^2=n$. Thus, $\mathrm{O}(n)$ is contained in the ball of radius $\sqrt{n}$ of $\mathrm{L}(n)$ and is bounded. ■

10. The special orthogonal group $\mathrm{SO}(n)=\mathrm{SO}_+(n)$ consists of the orthogonal matrices with determinant $+1$. It is a subgroup of $\mathrm{O}(n)$ and an open subset of it (why?). Hence, it is a Lie group of dimension equal to that of $\mathrm{O}(n)$.

11. The unitary group $\mathrm{U}(n)$. An argument similar to that used for $\mathrm{O}(n)$ shows that the unitary group $\mathrm{U}(n):=\big\{A\in \mathrm{GL}(n,\mathbb{C})\ :\ AA^*=\mathbb{I}\big\}$ is a (real) submanifold of $\mathrm{L}(n,\mathbb{C})$ (thought of as a $2n^2$-dimensional real vector space) whose codimension equals the dimension of the subspace of $\mathrm{L}(n,\mathbb{C})$ that consists of all hermitian matrices (thought of as a real vector space). Since in an $n\times n$ hermitian matrix one can freely choose the n diagonal entries, which are real, and the real and imaginary parts of the $\frac{1}{2}n(n-1)$ entries below the diagonal, this subspace has dimension n^2. Therefore, $\mathrm{U}(n)$ is a Lie group of dimension n^2.

Proposition 1.1.3 *For any* $n\geq 1$, $\mathrm{U}(n)$ *is compact and connected.*

Proof. Compactness is proved as for $\mathrm{O}(n)$. Next, recall that any unitary matrix A is (unitarily) similar to a diagonal matrix: there exists an invertible (actually, unitary) matrix P and a diagonal matrix D such that $A=PDP^{-1}$. On the other hand, the eigenvalues of a unitary matrix are complex numbers of modulus one. Hence

$$A=P\mathrm{diag}\left(e^{i\theta_1},\dots,e^{i\theta_n}\right)P^{-1}$$

with $\theta_1,\dots,\theta_n\in\mathbb{R}$. The path

$$[0,1]\ni t\mapsto P\mathrm{diag}\left(e^{it\theta_1},\dots,e^{it\theta_n}\right)P^{-1}$$

joins the identity to A. Thus, $\mathrm{U}(n)$ is pathwise-connected and hence connected. ■

12. The special unitary group. $\mathrm{SU}(n):=\{A\in\mathrm{U}(n)\ :\ \det A=1\}$ is a Lie group of dimension n^2-1. To prove this, first observe that the determinant of a unitary matrix is a complex number of absolute value 1 because, if $AA^*=\mathbb{I}$, then $|\det A|^2 = |\det(AA^*)| = 1$. Clearly, $\det : \mathrm{U}(n) \to S^1$ is surjective (see the Exercises).

Thus, let us regard the determinant as the map

$$\det : \mathrm{U}(n) \to S^1 ,$$

which is a smooth map between smooth manifolds. If we prove that this map has rank 1 at each point $A \in \mathrm{SU}(n)$, then we conclude that $\mathrm{SU}(n) = \det^{-1}(1)$ is a codimension 1 submanifold of $\mathrm{U}(n)$.

Saying that det has rank 1 at a point $A \in \mathrm{SU}(n)$ means that there exists a vector $V \in T_A\mathrm{U}(n)$ such that $T_A \det \cdot V \neq 0$. Fix $A \in \mathrm{SU}(n)$ and consider the curve

$$\gamma : \mathbb{R} \to \mathrm{U}(n) , \qquad t \mapsto e^{it}A$$

which passes through A at $t = 0$. Its derivative $\gamma'(0) = iA$ is a vector in $T_A\mathrm{U}(n)$ and

$$T_A \det \cdot iA \;=\; \tfrac{d}{dt} \det(e^{it}A)|_{t=0} \,.$$

But, $\det(e^{it}A) = e^{int}\det(A) = e^{int}$ (since $A \in \mathrm{SU}(n)$) and hence $T_A \det \cdot iA = ni \neq 0$.

$\mathrm{SU}(n)$ is compact (being a closed subset of $\mathrm{U}(n)$) and connected.

13. S^3. It turns out that, among all spheres $S^n := \{x \in \mathbb{R}^{n+1} \,:\, \|x\| = 1\}$, only two can be given the structure of a Lie group: S^1 and S^3. We describe here the Lie group structure of S^3. As a manifold, it has dimension 3; in all computations below, we will tacitly embed it in $\mathbb{R}^4$. The group structure of S^3 comes from the restriction to S^3 of the well known quaternion product of $\mathbb{R}^4$.

Recall that an *algebra* $\mathcal{A}$ is a vector space together with a bilinear operation $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$, called the 'product'. Denote $x = (x_0, \overline{x})$, with $\overline{x} = (x_1, x_2, x_3)$, the points of $\mathbb{R}^4$ (occasionally, we will say that x_0 and $\overline{x}$ are, respectively, the 'scalar' and 'vector' parts of $x = (x_0, \overline{x})$). Then, the quaternion product in $\mathbb{R}^4$ is defined as

$$(x_0, \overline{x})(y_0, \overline{y}) \;:=\; \big(x_0y_0 - \overline{x}\cdot\overline{y}\,,\; x_0\overline{y} + y_0\overline{x} + \overline{x}\times\overline{y}\big) \tag{1.1.3}$$

where $\cdot$ and $\times$ are the standard inner and cross products in $\mathbb{R}^3$. This gives $\mathbb{R}^4$ the structure of an algebra, the quaternion algebra. As is easily checked, this algebra is associative, has a unity (the element $(1, \overline{0})$) and is a division algebra (every nonzero element has an inverse—which one?), but is not abelian. Since the null vector 0 does not have an inverse $\mathbb{R}^4$ with this product is not a group.

However, a computation shows that, denoting $\|\ \|$ the Euclidean norm in $\mathbb{R}^4$, one has $\|xy\| = \|x\|\|y\|$ for all $x, y \in \mathbb{R}^4$. Thus, the quaternion product (1.1.3) restricts to a product on the unit sphere $S^3 \subset \mathbb{R}^4$, and gives it a group structure, with group identity

$$e_{S^3} = (1, \overline{0}) \,.$$

Since the product (1.1.3) and the associated inverse are smooth in $\mathbb{R}^4$ and therefore on its submanifold S^3, S^3 with the quaternion product (1.1.3) is a Lie group of dimension 3. Obviously, it is compact and connected.

Exercises 1.1.2 *(i) Verify that the Frobenius inner product in* $\mathrm{L}(n)$ *becomes, under the identification of* $\mathrm{L}(n)$ *and* $\mathbb{R}^{n^2}$ *via the basis* e_{ij} *as in (1.1.1), the Euclidean inner product in* $\mathbb{R}^{n^2}$.

(ii) Show that $\det : \mathrm{U}(n) \to S^1$ *is surjective. [Suggestion: consider, e.g., the matrices of the form* $e^{i\theta}\mathbb{I}$, $\theta \in \mathbb{R}$*].*

(iii) Show that the inverse of $(x_0, \overline{x}) \in S^3$ *is* $(x_0, -\overline{x})$.

(iv) Let $\mathcal{Q}$ *be the quaternion algebra.* $\mathcal{Q} \setminus \{0\}$ *is a group?*

(v) Often, the quaternion algebra is defined in a different way. Recall that a linear map $f : \mathcal{A} \to \mathcal{B}$ *between two algebras* $\mathcal{A}$ *and* $\mathcal{B}$ *is an algebra homomorphism if* $f(a, a') = f(a)f(a')$ *for all* $a, a' \in \mathcal{A}$*. Consider a real 4-dimensional vector space* $\mathcal{Q}$ *with a basis formed by 4 vectors, conventionally denoted* $1, i, j, k$*. Define a product on* $\mathcal{Q}$ *extending by linearity the relations*

$$1^2 = 1\,, \qquad i^2 = j^2 = k^2 = ijk = -1 \tag{1.1.4}$$

(from which all other products among the 4 basis vectors follow). Verify that the linear map $\mathcal{Q} \to \mathbb{R}^4$ *defined by* $1 \mapsto (1, \overline{0})$, $i \mapsto (0, \overline{e}_1)$, $j \mapsto (0, \overline{e}_2)$, $k \mapsto (0, \overline{e}_3)$ *is an isomorphism between* $\mathcal{Q}$ *with this product and* $\mathbb{R}^4$ *with the product (1.1.3). (Here* e_1, e_2, e_3 *are the vectors of the canonical basis of* $\mathbb{R}^3$*).*

(vi) Verify that a 2×2 *complex matrix belongs to* $\mathrm{SU}(2)$ *if and only if it has the form* $\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}$ *with* $a, b \in \mathbb{C}$ *such that* $|a|^2 + |b|^2 = 1$.

1.1.C Left and right translations.

Definition 1.1.4 *Let* G *be a Lie group and* $g \in G$*. The* left translation *by* g *is the map*

$$L_g \,:\, G \to G\,, \qquad h \mapsto gh\,,$$

the right translation *by* g *is the map*

$$R_g \,:\, G \to G\,, \qquad h \mapsto hg$$

and the conjugation *by* g *is the map*

$$C_g = L_g \circ R_{g^{-1}} \,:\, G \to G\,, \qquad h \mapsto ghg^{-1}\,.$$

Proposition 1.1.5 *For any* $g \in G$, L_g, R_g *and* C_g *are diffeomorphisms of* G *onto itself.*

Proof. L_g is smooth and is invertible, with inverse $L_{g^{-1}}$ which is smooth as well. Similarly for R_g and C_g. ∎

1.1.D Lie group homomorphisms and isomorphisms.

Definition 1.1.6

i. A Lie group homomorphism *between two Lie groups G and H is a group homomorphism between G and H which is also a smooth map.*

i. A Lie group isomorphism *is a Lie group homomorphism which is also a diffeomorphism.*

Examples: 1. The real exponential $\exp : \mathbb{R} \to \mathbb{R}^*$, $t \mapsto e^t$, is a Lie group homomorphism between $(\mathbb{R}, +)$ and $(\mathbb{R}^*, x)$. It is not an isomorphism because it is not surjective.

2. $\det : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}^*$ is a Lie group homomorphism.

3. The map

$$\varphi : \mathbb{T}^1 \to S^1 , \qquad x \mapsto e^{ix}$$

is a Lie group isomorphism.

4. Consider two reals ω_1 and ω_2. The map

$$\gamma : \mathbb{R} \to \mathbb{T}^2 , \qquad t \mapsto (\omega_1 t, \omega_2 t) \,\mathrm{mod} 2\pi \tag{1.1.5}$$

is a homomorphism of Lie groups ($\gamma(x+y) = \gamma(x)+\gamma(y) \,\mathrm{mod} 2\pi$ for all $x, y \in \mathbb{R}$). It is well known (Krönecker's theorem) that $\gamma(\mathbb{R})$ is a closed curve if ω_1 and ω_2 are linearly dependent over the rationals, and if not, it is dense in $\mathbb{T}^2$..

5. For any $g \in G$, the conjugation C_g is a Lie group isomorphism of G into itself (it is a diffeomorphism and $C_g(hh') = C_g(h)C_g(h')$ for all $h, h' \in G$).

Lie group homomorphisms are a generalization of linear maps between vector spaces. Linear maps $f : x \mapsto Ax$ have special properties: their fibers differ by translations ($f^{-1}(y) = f^{-1}(0) + v$ with any $v \in A^{-1}(y)$) and their derivative (being constant[2]) have constant rank.

Proposition 1.1.7 *If $f : G \to H$ is a Lie group homomorphism, then:*

i. The fiber of f that contains a point $g \in G$ is the left-translation by g of the fiber of f that contains e.

ii. f has constant rank.

Proof. (Recall that a group homomorphism $f : G \to H$ maps the identity element e of G into the identity element e_H of H.)

(i.) This is well known from group theory: the fibers of a group homomorphisms are the cosets of its kernel:

$$f^{-1}(f(g)) = gf^{-1}(e_H) \qquad \forall g \in G\,.$$

[2]This is meaningful because the tangent spaces to a vector space can be identified with the vector space itself.

Let us anyway detail it. If $g_1 \in f^{-1}(e_H)$, or $f(g_1) = e_H$, then $f(L_g g_1) = f(gg_1) = f(g)e_H = f(g)$ and so $L_g g_1 \in f^{-1}(f(g))$. Thus $L_g(f^{-1}(e_H)) \subseteq f^{-1}(f(g))$. Conversely, assume $g_2 \in f^{-1}(f(g))$. Thus $f(g^{-1}g_2) = f(g^{-1})f(g_2) = f(g)^{-1}f(g) = e_H$ and $g^{-1}g_2 \in f^{-1}(e_h)$. Hence $L_{g^{-1}}(f^{-1}(f(g))) \subseteq f^{-1}(e_H)$ and, since L_g is a diffeomorphism, $f^{-1}(f(g)) \subseteq L_g(f^{-1}(e_H))$.

(ii.) The condition that f is a homomorphism ($f(gg') = f(g)f(g')$ or $f(L_g g') = L_{f(g)}g'$ for all $g, g' \in G$) can be written

$$f \circ L_g = L_{f(g)} \circ f \qquad \forall g \,.$$

Using the chain rule, compute

$$\begin{aligned} T_e(f \circ L_g) &= T_{L_g e} f \circ T_e L_g = T_g f \circ T_e L_g \\ T_e(L_{f(g)} \circ f) &= T_{f(e)} L_{f(g)} \circ T_e f = T_{e_H} L_{f(g)} \circ T_e f \,. \end{aligned}$$

Since left translations are diffeomorphisms, the linear maps $T_e L_g$ and $T_{e_H} L_{f(g)}$ are isomorphisms. Thus, this gives

$$T_g f = T_{e_H} L_{f(g)} \circ T_e f \circ (T_e L_g)^{-1}$$

which implies that the linear maps $T_g f$ e $T_e f$ have the same rank. ■

1.1.E Lie subgroups. So far, we have been somehow unprecise about submanifolds. A subset S of a manifold M is said to be an *immersed submanifold* of M if there are a manifold $\tilde{S}$ and an injective immersion $j : \tilde{S} \to M$ such that $j(\tilde{S}) = S$. Sometimes, and in fact more precisely, the immersed submanifold is defined as the pair $(\tilde{S}, j)$. If $(\tilde{S}, j)$ is an immersed submanifold, then $j(\tilde{S})$ has a (unique) manifold structure such that $j : \tilde{S} \to S = j(\tilde{S})$ is a diffeomorphism.[3]

An immersed submanifold $(\tilde{S}, j)$ is said to be an *embedded submanifold*, or simply a *submanifold*, if the manifold topology of $j(\tilde{S})$ is the induced topology.[4] This is equivalent to the fact that $j : \tilde{S} \to M$ is an embedding[5]. Proper[6] injective immersions are embeddings. In particular, any compact immersed submanifold is an embedded submanifold.

Definition 1.1.8 *A subset H of a Lie group G is a* Lie subgroup *of G if it is a subgroup and an immersed submanifold of G.*

[3]This manifold structure is given by the final topology ($U \subset S$ is open if and only if $j^{-1}(S)$ is open in $\tilde{S}$) and by an atlas with charts of the form $\phi \circ (j|_S)^{-1}$, where the ϕ's are charts of $\tilde{S}$).

[4]Also called relative or subset topology: the open sets of $S = j(\tilde{S})$ are the intersection of S and the open sets of M.

[5]Embedding=injective immersion which is a homeomorphism onto its image.

[6]A map is proper if the preimages of compact sets are compact.

Equivalently, H is a Lie subgroup of G if there are a Lie group $\tilde{H}$ and an injective Lie group homomorphism $j : \tilde{H} \to G$ such that $H = j(\tilde{H})$. Hence, H is isomorphic, as a Lie group, to $\tilde{H}$.

> **Examples:** 1. All subgroups of $\mathrm{GL}(n)$ introduced in the previous section are Lie subgroups of $\mathrm{GL}(n)$ (in fact, they are embedded submanifolds).
>
> 2. For any $k \in \mathbb{Z}_+$, the map $\gamma : \mathbb{T}^1 \to \mathbb{T}^2$ given by $t \mapsto (t, kt) \bmod 2\pi$ is a Lie group isomorphism and thus $\gamma(\mathbb{T}^1)$ is a one-dimensional compact Lie subgroup of $\mathbb{T}^2$, isomorphic to $\mathbb{T}^1$.
>
> 3. If $\omega_1/\omega_2 \notin \mathbb{Q}$, then the image $\gamma(\mathbb{R})$ of the map (1.1.5) is a non-compact Lie subgroup of $\mathbb{T}^2$ isomorphic to $\mathbb{R}$.

In general, subgroups of a Lie group are not necessarily Lie subgroups. However, we state without proof the following result:

Proposition 1.1.9 *Any subgroup of a Lie group which is a closed subset is an embedded Lie subgroup.*

The closed subgroups of $\mathrm{GL}(n)$ are called *classical groups*, or *linear groups* or *matrix Lie groups*. All subgroups of $\mathrm{GL}(n)$ considered in the examples of section 1.1.B are of this type.

1.1.F S^3 and SU(2). As a final example, we investigate the relationships between the three 3-dimensional, compact and connected Lie groups we have met: S^3, $SU(2)$ and SO(3). We begin from the first two.

Refer to exercise 1.1.2.v. Consider the algebra $\mathbb{R}^4$ with the quaternion product (1.1.3) and the algebra $\mathrm{L}(2,\mathbb{C})$ with the matrix product. The linear map $f : \mathbb{R}^4 \to \mathrm{L}(2,\mathbb{C})$ given by

$$f \;:\; (x_0, \overline{x}) = (x_0, x_1, x_2, x_3) \mapsto \begin{pmatrix} x_0 + ix_1 & x_2 + ix_3 \\ -x_2 + ix_3 & x_0 - ix_1 \end{pmatrix}$$

is injective.

If $e_i = (1,0,0)$, $e_j = (0,1,0)$ and $e_k = (0,0,1)$, then the four matrices

$$\begin{aligned} M_1 &:= f(1,\overline{0}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & M_i &:= f(0,\overline{e}_i) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \\ M_j &:= f(0,\overline{e}_j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, & M_k &:= f(0,\overline{e}_k) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \end{aligned}$$

satisfy $M_1^2 = M_1$, $M_i^2 = M_j^2 = M_k^2 = M_i M_j M_k = -M_1$. Hence, a glance at formulas (1.1.4) shows that f is an algebra isomorphism onto its image.

Since the products of S^3 and SU(2) are the restrictions of the products of the two algebras $\mathbb{R}^4$ e $\mathrm{L}(2,\mathbb{C})$, this implies that $f|_{S^3} : S^3 \to \mathrm{SU}(2)$ is a group isomorphism. On the other hand, recalling from Exercise 1.1.2.vi the structure of the matrices in SU(2), we see that the restriction of f to the unit sphere S^3 of $\mathbb{R}^4$ is a bijective map $S^3 \to \mathrm{SU}(2)$ and hence, since $f : \mathbb{R}^4 \to f(\mathbb{R}^4) \subset \mathrm{L}(2, \mathbb{C})$ is a diffeomorphism, is a diffeomorphism.[7]

We conclude that $f|_{S^3} : S^3 \to \mathrm{SU}(2)$ is a Lie group isomorphism.

1.1.G S^3 and SO(3). We now compare S^3 and SO(3). In so doing, we will encounter some objects which, as will become clear later, have a Lie-theoretical nature.

1. The isomorphism of vector spaces $\wedge : \mathbb{R}^3 \to \mathrm{skew}(3)$. The set skew(3) of all 3×3 antisymmetric real matrices is a 3-dimensional linear subspace of the vector space L(3). The vector product with a fixed vector $u \in \mathbb{R}^3$ is a linear antisymmetric map $v \mapsto u \times v$ and is therefore represented by a (unique) matrix $\hat{u} \in \mathrm{skew}(3)$. This defines a map $\hat{} : \mathbb{R}^3 \to \mathrm{skew}(3)$ such that $\hat{u}v = u \times v$ for all $u, v \in \mathbb{R}^3$, which is given by

$$u \;=\; \begin{pmatrix} v_1 \\ u_2 \\ u_3 \end{pmatrix} \;\mapsto\; \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} . \tag{1.1.6}$$

Clearly, this map is a linear isomorphism between $\mathbb{R}^3$ and skew(3).

We note the following two properties of the isomorphism $\hat{}$:

$$\widehat{u \times v} \;=\; \hat{u}\hat{v} - \hat{v}\hat{u} \qquad \forall\, u, v \in \mathbb{R}^3 \tag{1.1.7}$$

(which, at this stage, can be verified with a non-enlightening direct computation) and

$$\widehat{Ru} \;=\; R\hat{u}R^T \qquad \forall\, R \in \mathrm{SO}(3) . \tag{1.1.8}$$

(in fact, $\widehat{Ru}\,v = (Ru) \times v = R(u \times R^T v) = R(\hat{u}R^T v)$).

2. The matrix exponential. Let $\exp : \mathrm{L}(n) \to \mathrm{L}(n)$ be the matrix exponential map, which is defined as the sum of the series

$$\exp(A) := \sum_{n=0}^{\infty} \frac{1}{n!} A^n . \tag{1.1.9}$$

This series is uniformly convergent on compact sets, and exp is thus a smooth map. Obviously, $\exp(\hat{0}) = \mathbb{I}$. Moreover, the matrix exponential map has the following (well known, or easy to prove) properties:

Proposition 1.1.10 *For all $A \in \mathrm{L}(n)$:*

[7] The restriction of a diffeomorphism to an embedded submanifold is a diffeomorphism.

i. $\exp(A) \in \mathrm{GL(n)}$ *and* $\exp(A)^{-1} = \exp(-A)$.
ii. $\exp(A)^T = \exp(A^T)$.
iii. $P^{-1}\exp(A)P = \exp(P^{-1}AP)$ *for all* $P \in \mathrm{GL(n)}$
iv. $\det(\exp(A)) = e^{\mathrm{tr}\,(A)}$
v. $\frac{d}{dt}\exp(At) = A\exp(At) = \exp(At)A$ *for all* $t \in \mathbb{R}$

At the moment, we are particularly interested in the exponential of *skew-symmetric* matrices.

Proposition 1.1.11
i. *If* $A \in \mathrm{skew}(3)$ *then* $\exp(A) \in \mathrm{SO}(3)$.
ii. $\exp(\hat\omega) = \mathbb{I} + \frac{\sin\|\omega\|}{\|\omega\|}\,\hat\omega + \frac{1-\cos\|\omega\|}{\|\omega\|^2}\,\hat\omega^2$ *for all* $\omega \in \mathbb{R}^3\setminus\{0\}$ *("Euler-Rodrigues formula").*
iii. $\exp|_{\mathrm{skew}(3)} : \mathrm{skew}(3) \to \mathrm{SO}(3)$ *is surjective.*

Proof. (i) If $A = -A^T$ then $\exp(A)^T = \exp(A^T) = \exp(-A) = \exp(A)^{-1}$ and $\exp(A) \in \mathrm{O}(3)$. But $\mathrm{tr}\,(A) = 0$ and so $\det(\exp(A)) = e^0 = 1$.

(ii) Use identities (1.1.11) and rearrange the terms in the exponential series, collecting the terms linear and quadratic in $\hat\omega$.

(iii) We first prove that, if $R \in \mathrm{SO}(3)$ fixes a vector $\omega \in \mathbb{R}^3$, namely $R\omega = \omega$, then $R = \exp(k\hat\omega)$ with a certain $k \in \mathbb{R}$.

Preliminarily, assume that $P \in \mathrm{SO}(3)$ fixes $e_3 = (0,0,1)$. Then $P = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$ with some $\alpha \in \mathbb{R}$. Since $\hat{e_3} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, and thus $\hat{e_3}^2 = -\mathbb{I} + e_3e_3^T$, from the Euler-Rodrigues formula it follows that $P = \exp(\alpha\hat{e_3})$. Consider now an $\omega \in \mathbb{R}^3 \setminus \{0\}$. Choose a matrix $S \in \mathrm{SO}(3)$ such that $S\omega = \|\omega\|e_3$. Then $SRS^Te_3 = e_3$ and so $SRS^T = \exp(\alpha\hat{e_3})$ for some α. It follows that $R = S^{-1}\exp(\alpha\hat{e_3})S = \exp(\alpha S^{-1}\hat{e_3}S) = \exp(\alpha S^T\hat{e_3}S) = \exp(\alpha\widehat{S^Te_3}) = \exp\left(\frac{\alpha}{\|\omega\|}\hat\omega\right)$.

In view of this, in order to prove the surjectivity we just need to prove that any matrix $R \in \mathrm{SO}(3)$ has the eigenvalue 1. All eigenvalues of a matrix $R \in \mathrm{SO}(3)$ have modulus one (if $Rx = \lambda x$ with $\lambda \in \mathbb{C}$ and $x \neq 0$ then $\|x\| = \|Rx\| = |\lambda|\|x\|$ and hence $|\lambda| = 1$). Together with $\det R = 1$ this implies that R has the eigenvalue 1. (If R has one real eigenvalue λ and a pair of complex conjugate eigenvalues α and $\overline{\alpha} \neq \alpha$ then $1 = \det R = |\alpha|^2\lambda$ and $\lambda = +1$. If R has three real eigenvalues, they cannot all be -1 because their product is $+1$). ■

3. A Lie group homomorphism $S^3 \to \mathrm{SO}(3)$. In the sequel, if $q = (q_0, \overline{q}) \in \mathbb{R}^4$, we write $\hat q$ for $\widehat{\overline{q}}$. For all $q = (q_0, \overline{q}) \in \mathbb{R}^4$ define

$$\mathcal{E}(q) := \mathbb{I} + 2q_0\hat q + 2\hat q^2\,. \tag{1.1.10}$$

Proposition 1.1.12 *The restriction of $\mathcal{E}$ to S^3 is a Lie group homomorphism between S^3 and* $\mathrm{SO}(3)$.

Proof. First, we show that $\mathcal{E}(q) \in \mathrm{SO}(3)$ for all $q \in S^3$. Indeed, for all $q \in \mathbb{R}^4$,

$$\mathcal{E}(q)\mathcal{E}(q)^T \;=\; \mathbb{I} + 4\big((1-q_0^2)\hat{q}^2 + \hat{q}^4\big) \;=\; \mathbb{I} + 4\big(1 - q_0^2 - \|\overline{q}\|^2\big)\hat{q}^2$$

where the last expression follows from $\hat{q}^4 = -\|\overline{q}\|^2\hat{q}^2$ (see Exercise 1.1.3.i). If $q \in S^3$ then $q_0^2 + \|\overline{q}\|^2 = 1$ and $\mathcal{E}(q)\mathcal{E}(q)^T = \mathbb{I}$. Thus $\mathcal{E}(S^3) \subseteq \mathrm{O}(3)$. Since S^3 is connected and $\mathcal{E}$ is continuous, $\mathcal{E}(S^3)$ is connected. And since $\mathcal{E}(1,0) = \mathbb{I} \in \mathrm{SO}(3)$, $\mathcal{E}(S^3) \subseteq \mathrm{SO}(3)$.

Next, smoothness of $\mathcal{E}$ is obvious and a computation shows that $\mathcal{E}(qp) = \mathcal{E}(q)\mathcal{E}(p)$ for all $q, p \in S^3$ (see Exercise 1.1.3.ii). ■

Proposition 1.1.13 *The map* $\mathcal{E} : S^3 \to \mathrm{SO}(3)$*:*

i. Is $2:1$.

ii. Is surjective

iii. Is a local diffeomorphism.

Proof. (i) Clearly $\mathcal{E}(q) = \mathcal{E}(-q)$ for all $q \in S^3$ and therefore each fiber of $\mathcal{E}$ contains at least two points. But it is easily seen that the fiber $\mathcal{E}^{-1}(\mathbb{I})$ consists of exactly two points. Indeed $\mathcal{E}(q) = \mathbb{I} + 2q_0\hat{q} + 2\hat{q}^2 = \mathbb{I}$ if and only if $q_0\hat{q} = -\hat{q}^2$, which implies $\hat{q} = 0$ because $\hat{q}$ is antisymmetric and $\hat{q}^2$ is symmetric. Hence $q_0 = \pm 1$ and $\mathcal{E}^{-1}(\mathbb{I}) = \{(1,\overline{0}), (-1,\overline{0})\}$. Hence, by Proposition 1.1.7, all other fibers of $\mathcal{E}$ have cardinality 2.

(ii) Since $\exp \circ \,\hat{}\, : \mathbb{R}^3 \to \mathrm{SO}(3)$ is surjective, to prove the surjectivity of $\mathcal{E} : S^3 \to \mathrm{SO}(3)$ we show that, for any $\omega \in \mathbb{R}^3$, there exists a $q = (q_0, \overline{q}) \in S^3$ such that $\mathcal{E}(q) = \exp\hat{\omega}$. Being a homomorphism, $\mathcal{E}$ maps the identity element $(1, \overline{0})$ of S^3 into the identity element $\mathbb{I} = \exp 0$ of $\mathrm{SO}(3)$. So, we may limit ourselves to $\omega \neq 0$. With a little trigonometry, the Euler-Rodrigues formula may be rewritten as

$$\exp\hat{\omega} \;=\; \mathbb{I} + 2\left(\sin\frac{\|\omega\|}{2}\right)\left(\cos\frac{\|\omega\|}{2}\right)\frac{\hat{\omega}}{\|\omega\|} + 2\left(\sin\frac{\|\omega\|}{2}\right)^2\frac{\hat{\omega}^2}{\|\omega\|^2}\,.$$

This equals $\mathbb{I} + 2q_0\hat{q} + 2\hat{q}^2$ if $q_0 = \cos\frac{\|\omega\|}{2}$ and $\overline{u} = \frac{\overline{\omega}}{\|\omega\|}\sin\frac{\|\omega\|}{2}$; the resulting point $q = (q_0, \overline{q}) \in \mathbb{R}^4$ belongs to S^3 because $q_0^2 + \|\overline{q}\|^2 = 1$.

(iii) This is equivalent to prove that $\mathcal{E}$ has rank 3 at all points of S^3. By Proposition 1.1.7, it is sufficient to verify this fact at the identity $(1, \overline{0})$ of S^3.

Embed S^3 and its tangent spaces in $\mathbb{R}^4$. For any $q = (q_0, \overline{q}) \in S^3$,

$$\begin{aligned} T_qS^3 \;&=\; \{v \in \mathbb{R}^4 \;:\; q \cdot v = 0\} \\ &=\; \{(v_0, \overline{v}) \in \mathbb{R} \times \mathbb{R}^3 \;:\; q_0 v_0 + \overline{q} \cdot \overline{v} = 0\}\,. \end{aligned}$$

In particular, $T_{(1,\overline{0})}S^3 = \{(0,\overline{v}) \;:\; \overline{v} \in \mathbb{R}^3\}$. If $q \in S^3$ e $v \in T_qS^3$ then $v = \frac{d}{dt}(q+tv)|_{t=0}$ and

$$\begin{aligned} T_q\mathcal{E}\cdot v &= \tfrac{d}{dt}\mathcal{E}(q+tv)|_{t=0} \\ &= \tfrac{d}{dt}\left[\mathbb{I} + 2t(q_0\hat{v} + v_0\hat{q} + 2\hat{q}\hat{v}) + \mathcal{O}(t^2)\right]_{t=0} \\ &= 2(q_0\hat{v} + v_0\hat{q} + 2\hat{v}\hat{q})\,. \end{aligned}$$

Thus $T_{(1,\overline{0})}\mathcal{E}\cdot v = 2\hat{v}$ and so $T_{(1,\overline{0})}$ is the map $(0,\overline{v}) \mapsto 2\hat{v}$, which has rank 3. ■

Thus S^3, and hence $SU(2)$, is not isomorphic to SO(3), but it is homomorphic and a double covering of SO(3). [8]

> **Remarks:** (i) The fact that every rotation matrix has the eigenvalue +1 is sometimes called 'Euler theorem on the rigid body motion'. In Lie-group terms, the surjectivity of the map exp : skew(3) → SO(3), known to Euler, means that the exponential map of SO(3) is surjective (as, we will see, for any compact and connected Lie group, but not in general).
>
> (ii) The Euler parameters (or Euler-Rodrigues parameters) map $\mathcal{E}$ provides a parametrization of 3d rotations by means of unit quaternions. Even though this parametrization is not unique—there are two quaternions that correspond to each rotation—there are various other reasons that make this parametrization very useful and commonly used in mechanics, engineering and computer graphics. First, this parametrization allows an immediate geometric comprehension of the rotation: $\mathcal{E}(q)$ is the rotation of axis $\overline{q}$ ed angle α such that $\cos(\alpha/2) = q_0$ and $\sin(\alpha/2) = \|\overline{q}\|$. Second, the fact that $\mathcal{E}(q)\mathcal{E}(p) = \mathcal{E}(qp)$ makes easy to determine the axis of the composition of two rotations. And furthermore, even though it uses 4 parameters instead of 3 coordinates, it has the advantage of providing a global (though not 1:1) parametrization, while every atlas of SO(3) requires at least two charts (with complicated transition functions).

Exercises 1.1.3 *(i) Prove that*

$$\hat{\omega}^{2n+1} = (-1)^n\|\omega\|^{2n}\,\hat{\omega}\,, \qquad \hat{\omega}^{2n+2} = (-1)^n\|\omega\|^{2n}\,\hat{\omega}^2\,, \qquad n \in \mathbb{N} \tag{1.1.11}$$

[Hints: A simple computation (or a picture) gives $\hat{\omega}^2 = \|\omega\|^2(-\mathbb{I}+\Pi_\omega)$, with Π_ω the orthogonal projection onto the subspace spanned by ω. Thus $\hat{\omega}^3 = -\|\omega\|^2\hat{\omega}$, $\hat{\omega}^4 = -\|\omega\|^2\hat{\omega}^2$. Use induction].

---

[8] Let M and P be two connected manifolds. A submersion $\pi : M \to P$ is said to be a *covering* if, for each $p \in P$, there exist a neighbourhood $U \subset P$ of p and, for each $m \in \pi^{-1}(p)$, a neighbourhood $V \subset M$ of m such that $\pi|_V$ is a diffeomorphism from V onto U. If there exists $k \in \mathbb{N}$ such that each point $p \in P$ has k preimages, then the covering is said to have k *sheets.* A basic example of covering is the map $\pi : \mathbb{R} \to S^1 \subset \mathbb{C}$, $\pi(t) = e^{it}$. A local diffeomorphism which is surjective and whose fibers have constant cardinality $k \in \mathbb{Z}_+$ is a k-sheeted covering.

(ii) Verify that $\mathcal{E}(uv) = \mathcal{E}(u)\mathcal{E}(v)$ *for all* $u, v \in S^3$. *[Hints:* $\mathcal{E}(uv) = \mathbb{I} + 2(u_0 v_0 - \overline{u}\cdot\overline{v})(u_0\hat{v} + v_0\hat{u} + \widehat{\overline{u}\times\overline{v}}) + 2(u_0\hat{v} + v_0\hat{u} + \widehat{\overline{u}\times\overline{v}})^2$, $\mathcal{E}(u)\mathcal{E}(v) = (\mathbb{I} + 2u_0\hat{u} + 2\hat{u}^2)(\mathbb{I} + 2v_0\hat{v} + 2\hat{v}^2)$ *and* $\widehat{\overline{u}\times\overline{v}} = \hat{u}\hat{v} - \hat{v}\hat{u}$.*]*

1.2 The Lie algebra of a Lie group

1.2.A Lie algebras

Definition 1.2.1

i. A Lie algebra *is an algebra* $\mathcal{A}$ *whose product* $[\ ,\ ] : \mathcal{A}\times\mathcal{A} \to \mathcal{A}$ *is antisymmetric and satisfies the* Jacobi identity

$$[u_1,[u_2,u_3]] + [u_2,[u_3,u_1]] + [u_3,[u_1,u_2]] \qquad \forall\, u_1,\, u_2,\, u_3 \in \mathcal{A}\,.$$

ii. If $(\mathcal{A}, [\ ,\ ])$ *is a Lie algebra, a linear subspace* $\mathcal{B}$ *of* $\mathcal{A}$ *is a* Lie subalgebra *of* $\mathcal{A}$ *if it is a Lie algebra with the (restriction of) the product* $[\ ,\ ]$ *of* $\mathcal{A}$.

iii. A Lie algebra homomorphism *between two Lie algebras* $(\mathcal{A}, [\ ,\ ]_{\mathcal{A}})$ *and* $(\mathcal{B}, [\ ,\ ]_{\mathcal{B}})$ *is a linear map* $f : \mathcal{A} \to \mathcal{B}$ *which satisfies* $f([u,v]_{\mathcal{A}}) = [f(u), f(v)]_{\mathcal{B}}$ *for all* $u, v \in \mathcal{A}$. *If such an* f *is a linear isomorphism, then it is said to be a* Lie algebra isomorphism.

Examples:

1. The linear operators on a vector space, with the commutator $[\ ,\ ]_-$ as product. The commutator of linear operators is bilinear and antisymmetric, and a computation shows that it satisfies the Jacobi identity: for any three linear operators A, B, C, $[A,[B,C]_-]_- + [B,[C,A]_-]_- + [C,[A,B]_-]_- = A(BC - CB) - (BC - CB)A + B(CA - AC) - (CA - AC)B + C(AB - BA) - (AB - BA)C = 0$.

2. $\mathrm{L}(n,\mathbb{R})$ *and* $\mathrm{L}(n,\mathbb{C})$ *with the matrix commutator* are Lie algebras. This can be proven as in Example 1 (or even seen as a special case of it). We will see that the Lie algebras of the classical Lie groups are Lie subalgebras of $\mathrm{L}(n,\mathbb{R})$ or $\mathrm{L}(n,\mathbb{C})$.

3. $\mathrm{skew}\,(n)$ *with the matrix commutator* is a Lie subalgebra of $\mathrm{L}(n)$. This follows from the fact that the commutator of two antisymmetric matrices is antisymmetric (because $[A,B]_-^T = -[A^T, B^T]_-$) and from Exercise 1.2.1.i.

4. $(\mathbb{R}^3, \times)$ *is a Lie algebra isomorphic to* $(\mathrm{skew}(3), [\ ,\ ]_-)$. The cross product is bilinear and antisymmetric. The isomorphism of vector spaces $\hat{\ } : \mathbb{R}^3 \to \mathrm{skew}(3)$ satisfies $\widehat{u\times v} = [\hat{u}, \hat{v}]$ for all $u, v \in \mathbb{R}^3$. This implies that $(\mathbb{R}^3, \times)$ is a Lie algebra and $\hat{\ }$ is a Lie algebra isomorphism between $(\mathbb{R}^3, \times)$ e $(\mathrm{skew}(3), [\ ,\ ]_-)$ (see Exercise 1.2.1.ii).

5. The vector fields on a manifold, with the Lie bracket as product. $\mathcal{X}(M)$ is an (infinite dimensional) vector space. The Lie bracket is bilinear and antisymmetric. To prove that it satisfies the Jacobi identity we use the fact that a vector field $W \in \mathcal{X}(M)$ is zero if and only if the associated Lie derivative

$L_W : C^\infty(M) \to C^\infty(M)$ is zero. Let Jac $:= [X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]$. Since $L_{[X,Y]} = [L_X, L_Y]_-$ etc, we have

$$L_{\rm Jac} \;=\; [L_X,[L_Y,L_Z]_-]_- \;+\; [L_Y,[L_Z,L_X]_-]_- \;+\; [L_Z,[L_X,L_Y]_-]_-$$

and this vanishes for any choice of X, Y, Z because the Lie derivatives are linear operators (see Example 1.).

6. The Poisson bracket. Consider $\mathbb{R}^{2n} \ni (q,p) = (q_1,\ldots,q_n,p_1,\ldots,p_n)$. The Poisson bracket of two functions f, g is defined as

$$\{f,g\} \;:=\; \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}\,.$$

As is known from courses in Hamiltonian mechanics, it satisfies the Jacobi identity. Thus $(C^\infty(\mathbb{R}^{2n},\mathbb{R}), \{\ ,\ \})$ is an (infinite dimensional) Lie algebra.

7. Hamiltonian vector fields. Same setting as in Example 6. The Hamiltonian vector field $X_h \in \mathcal{X}(\mathbb{R}^{2n})$ of a function $h \in C^\infty(\mathbb{R}^{2n})$ is defined as

$$X_h \;=\; \sum_{i=1}^n \left(\frac{\partial h}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial h}{\partial q_i}\frac{\partial}{\partial p_i} \right)\,.$$

Hamiltonian vector fields form an (infinite dimensional, of course) vector subspace $\mathcal{X}_{\rm Ham}(\mathbb{R}^{2n})$ of $\mathcal{X}(\mathbb{R}^{2n})$. It is known that the Lie bracket of two of them satisfies

$$[X_f, X_g] = -X_{\{f,g\}}\,.$$

Thus $\mathcal{X}_{\rm Ham}(\mathbb{R}^{2n})$ is a Lie subalgebra of $\mathcal{X}(\mathbb{R}^{2n})$, and is anti-homomorphic ('anti': because of the minus sign) to $C^\infty(\mathbb{R}^{2n})$ with the Poisson bracket. (Homomorphic, not isomorphic: why?).

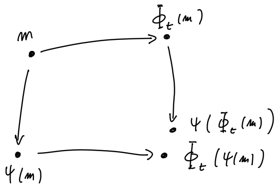
Exercises 1.2.1 *(i) Show that a vector subspace $\mathcal{B}$ of a Lie algebra $(\mathcal{A}, [\ ,\ ]_\mathcal{A})$ is a Lie subalgebra of $\mathcal{A}$ if and only if $[\mathcal{B},\mathcal{B}]_\mathcal{A} \subseteq \mathcal{B}$.*

(ii) Consider a Lie algebra $(\mathcal{A}, [\ ,\ ]_\mathcal{A})$, a vector space $\mathcal{B}$ and a linear isomorphism $f : \mathcal{A} \to \mathcal{B}$. Define the bilinear map $[\ ,\ ]_\mathcal{B}$ on $\mathcal{B}$ through

$$[f(u), f(v)]_\mathcal{B} := f([u,v]_\mathcal{A}) \qquad \forall\, u, v \in \mathcal{A}\,.$$

Show that $(\mathcal{B}, [\ ,\ ]_\mathcal{B})$ is a Lie algebra and that it is isomorphic to $(\mathcal{A}, [\ ,\ ]_\mathcal{A})$.

(iii) In Exercise (ii), under which additional assumptions may we replace 'linear isomorphism' with 'linear map' so as to get a Lie algebra structure on $\mathcal{B}$ (which is then only homomorphic to $\mathcal{A}$, of course)?



1.2.B Symmetries of a vector field. Consider a vector field X on a manifold M and a diffeomorphism Ψ of M onto itself. Consider the integral curve $t \mapsto \Phi_t^X(m)$ of X through a point $m \in M$ and its image $t \mapsto \Psi(\Phi_t^X(m))$ under Ψ. This latter curve passes through $\psi(m)$ but, in general, need not coincide with the integral curve of X through $\psi(m)$, namely $t \mapsto \Phi_t^X(\psi(m))$.

The symmetries of a vector field (or of the associated ODE) are those diffeomorphisms that map integral curves into integral curves, preserving their time parametrization, namely

$$\Psi \circ \Phi_t^X \;=\; \Phi_t^X \circ \Psi \qquad \forall\, t \in \mathbb{R}\,. \tag{1.2.1}$$

Since the vector field whose integral curves are the images under Ψ of those of X is $\Psi_* X$, this condition is equivalent to

$$\Psi_* X \;=\; X\,. \tag{1.2.2}$$

Recall that this condition is $T\Psi \cdot X = X \circ \Psi$ or, using local representatives, $\Psi' X = X \circ X$. The advantage of (1.2.2) over (1.2.1) is obviously due to the fact that, in general, the flow of a vector field cannot be determined.

Definition 1.2.2 *If (1.2.2) is satisfied, then the vector field X is said to be* invariant *under the diffeomorphism Ψ, and Ψ is said to be a* symmetry*, or a* symmetry transformation*, of X.*

Examples: 1. Consider the vector field $X(x, v) = (v, -x)$ on $\mathbb{R}^2 \ni (x, v)$ (a 'harmonic oscillator'). Let Ψ be the (anticlockwise) rotation of $\mathbb{R}^2$ of a certain angle α: $\Psi(z) = R_\alpha z$ with $R_\alpha \in SO(2)$. Since Ψ is linear, $\Psi' = R_\alpha$. Note now that

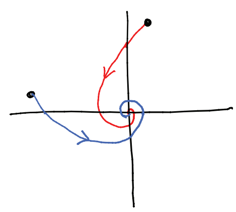
$$X(x, v) = \begin{pmatrix} v \\ -x \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix} = R_{-\pi/2} \begin{pmatrix} x \\ v \end{pmatrix}.$$

Hence, since planar rotations commute,

$$(\Psi' X)(x, v) = R_\alpha R_{-\pi/2} \begin{pmatrix} x \\ v \end{pmatrix} = R_{-\pi/2} R_\alpha \begin{pmatrix} x \\ v \end{pmatrix} = X\Big(\Psi(x, v)\Big) \quad \forall (x, v) \in \mathbb{R}^2$$

showing that Ψ is a symmetry of X. Since the flow of X consists of (clockwise) rotations, $\Psi_t^X = R_{-t}$, it is immediate to check that 1.2.1 is satisfied as well. Note that, since the integral curves of X are circles run at constant speed, in this case the symmetry Ψ maps each orbit (namely, the image of the integral curve) into itself. This is a special case.

2. Similarly, the planar rotation of any angle α is also a symmetry of the vector field $Y(x, y) = (-x - y, x - y)$ in $\mathbb{R}^2 \ni (x, y)$. In this case, however, the flow of Y is $\Phi_t^Y = e^{-t} R_t$ (verify it) and Ψ maps each orbit of X (except the equilibrium $(0, 0)$) into a different orbit of X.



Proposition 1.2.3 *The set of all vector fields on a manifold M which are invariant under a diffeomorphism $\Psi : M \to M$ is a Lie subalgebra of $\mathfrak{X}(M)$.*

Proof. Since condition (1.2.2) is linear in X, the set of Ψ-invariant vector fields is a vector subspace of $\mathfrak{X}(M)$. Since the Lie bracket is natural with respect to push-forward, namely $\Psi_*[X, Y] = [\Psi_* X, \Psi_* Y]$ for all $X, Y \in \mathfrak{X}(M)$, the Lie bracket of two Ψ-invariant vector field is Ψ-invariant. The conclusion follows from Exercise 1.2.1.i. ■

The same is true for the set of vector fields which are invariant under all diffeomorphisms of a given family of diffeomorphisms.

Exercises 1.2.2 *(i) Show that the linear map $x \mapsto Px$, where $P \in \mathrm{GL}(n)$, is a symmetry of a linear vector field $X(x) = Ax$ on $\mathbb{R}^n$, with $A \in \mathrm{L}(n)$, if and only if the two matrices A e P commute.*

(ii) Show that the set of all symmetries of a given vector field is a subgroup of the group of all diffeomorphisms of the manifold, with the composition as product.

(ii) Show that the set of all the equilibria of a vector field X is invariant under any symmetry of X (if $\overline{m}$ is an equilibrium, $\psi(\overline{m})$ is an equilibrium as well). Show that the same is true for the set of all periodic orbits of X and, actually, for the set of all periodic orbits of given period $T > 0$.

1.2.C Left-invariant vector fields on a Lie group.

This is a central notion in the theory:

Definition 1.2.4 *A vector field X on a Lie group G is* left-invariant *if it is invariant under all left-translations, namely*

$$(L_g)_* X \;=\; X \qquad \forall\, g \,\in\, G\,.$$

Similarly, X is right-invariant *if $(R_g)_* X = X$ for all $g \in G$.*

Examples: *1.* $\mathbb{R}^n$*.* Since $L_x y = x + y$, $(L_x)' = \mathbb{I}$ and, if V is a vector field in $\mathbb{R}^n$,

$$[(L_x)_* V](y) \;=\; (\mathbb{I}\, V)(y - x) \;=\; V(y - x)\,.$$

Thus, the left-invariance of V amounts to $V(y - x) = V(x)$ for all $x, y \in \mathbb{R}^n$, namely, to the constancy of V. This shows that left-invariant vector fields on a Lie group are a generalization of constant vector fields in $\mathbb{R}^n$. Note that constant vector fields in $\mathbb{R}^n$ are determined by their value at $0 \in \mathbb{R}^n$, which may be any vector in $\mathbb{R}^n$, and therefore form a vector space isomorphic to $\mathbb{R}^n$.

2. $\mathrm{GL}(n)$*.* Since $\mathrm{GL}(n)$ is an open submanifold of the vector space $\mathrm{L}(n)$, its tangent spaces may be identified with $\mathrm{L}(n)$, and the tangent bundle $T\mathrm{GL(n)}$ may be identified $\mathrm{GL(n)} \times \mathrm{L}(n)$. Neglecting the specification of the base point, a vector field X on $\mathrm{GL(n)}$ can be viewed as a map $A \mapsto X(A)$ from $\mathrm{GL(n)}$ to $\mathrm{L}(n)$.

Fix $A \in \mathrm{GL(n)}$. The left translation $L_A : \mathrm{GL(n)} \to \mathrm{GL(n)}$ is the restriction to $\mathrm{GL(n)}$ of the linear map $B \mapsto AB$ in $\mathrm{L}(n)$. Hence, for any $B \in \mathrm{GL(n)}$, the tangent map $T_B L_A : \mathrm{L}(n) \to \mathrm{L}(n)$ is the multiplication by A from the left: $T_B L_A \cdot V = AV$ for all $V \in \mathrm{L}(n)$. Thus

$$[(L_A)_* X](B) \;=\; (AX) \circ L_{A^{-1}} B \;=\; A\, X(A^{-1} B)$$

and equating this expression to $X(B)$ we conclude that X is left-invariant if and only if

$$X(B) \;=\; A\,X(A^{-1}B) \qquad \forall A, B \in \mathrm{GL}(n)\,. \tag{1.2.3}$$

Choosing $B = A$, this implies that, if X is left-invariant, then necessarily

$$X(A) \;=\; AX(\mathbb{I}) \qquad \forall A \in \mathrm{GL}(n)\,. \tag{1.2.4}$$

Conversely, for any $V \in \mathrm{L}(n)$, the vector field X_V on GL(n) defined by $X_V(A) = AV$ for all $A \in$ GL(n) is left-invariant because $T_B L_A \cdot X(B) = ABV = X(AB)$ $\forall B \in$ GL(n).

Hence, here too, left-invariant vector fields are determined by their value at the group identity and form a vector space isomorphic to $\mathrm{L}(n) = T_{\mathbb{I}}\mathrm{GL}(\mathrm{n})$. Moreover, they are obtained by transporting their value at $\mathbb{I}$ to all points of the group with (the tangent map of) left translations. As we now show, this is the case with all Lie groups and is at the basis of the construction of the Lie algebra of a Lie group.

From now on, we denote by $\mathfrak{X}_L(G)$ the vector space of all left-invariant vector fields on a Lie group G. Clearly, $\mathfrak{X}_L(G)$ is a vector subspace of $\mathfrak{X}(G)$ and, as noted at the end of Section 1.2.B, it is also a Lie subalgebra of the Lie algebra $(\mathfrak{X}(G), [\ ,\ ])$, where $[\ ,\ ]$ is the Lie bracket (or commutator[9]) of vector fields.

Proposition 1.2.5 *Let G be a Lie group. For any $\xi \in T_eG$, let X_ξ be the vector field on G defined by*

$$X_\xi(g) = T_e L_g \cdot \xi\,, \qquad g \in G\,.$$

Then:

i. $X_\xi \in \mathfrak{X}_L(G)$ *for all* $\xi \in T_eG$.

ii. The map

$$\lambda : T_eG \to \mathfrak{X}_L(G)\,, \quad \xi \mapsto X_\xi \tag{1.2.5}$$

is an isomorphism of vector spaces.

Proof. (i) First, note that since $T_eL_g \cdot \xi \in T_gG$ for all $g \in G$, X_ξ is actually a vector field. It is left-invariant because, for any $g, h \in G$, $[(L_g)_* X_\xi](h) = T_{g^{-1}h}L_g \cdot X_\xi(g^{-1}h) = T_{g^{-1}h}L_g \cdot T_eL_{g^{-1}h}\xi = T_e(L_g \circ L_{g^{-1}h}) \cdot \xi = T_eL_h \cdot \xi = X_\xi(h)$.

(ii) The map λ is clearly linear. It is injective because if $X_\xi = X_\eta$ for some $\xi, \eta \in T_eG$, then $\xi = X_\xi(e) = X_\eta(e) = \eta$. In order to show that is surjective we need showing that any left-invariant vector field X satisfies

$$X(g) \;=\; T_eL_g \cdot X(e)\,, \qquad g \in G\,,$$

[9] We will tend to use the term ‘commutator’ for the Le bracket of vector fields, and ‘Lie bracket’ for the product of a generic Lie algebra.

namely that $X = X_{X(e)}$. In fact, condition $(L_g)_* X = X$ can be written $TL_g \cdot X = X \circ L_g$. Evaluated at e, this gives $T_e L_g \cdot X(e) = X(L_g e) = X(g)$. ∎

As is known (see the Appendix), vector fields have in a way a double nature: they act as derivations and generate flows. Both aspects enter Lie groups: the first in the construction of the Lie algebra of a Lie group, and the second in the construction of the exponential map of a Lie group. We thus note here some properties of the flow of left-invariant vector fields. These vector fields generalize constant vector fields on $\mathbb{R}^n$, and have some of their properties: they are complete (namely, all solutions exist for all times) and their flow consists of (right) translations.

Proposition 1.2.6 *For any $\xi \in T_e G$:*

i. X_ξ *is complete.*

ii. $\Phi_t^{X_\xi}(g) = g\Phi_t^{X_\xi}(e)\ \forall\, g \in G$.

iii. $\Phi_{t+s}^{X_\xi}(e) = \Phi_t^{X_\xi}(e)\Phi_s^{X_\xi}(e) = \Phi_s^{X_\xi}(e)\Phi_t^{X_\xi}(e)\ \forall\, t, s \in \mathbb{R}$.

Proof. (i) The condition of left-invariance $(L_g)_* X = X\ \forall g \in G$ can be rewritten as $L_g \circ \Phi_t^X = \Phi_t^X \circ L_g$ for all $g \in G$ and those $t \in \mathbb{R}$ for which Φ_t^X is defined (see (1.2.1) and (1.2.2) with $\Psi = L_g$). Therefore, if $\gamma : I \to G$ is an integral curve of X through $e \in G$, then for any $g \in G$

$$L_g \circ \gamma \,:\, I \to G\,, \qquad t \mapsto L_g \gamma(t)\,,$$

is an integral curve of X through g. Hence, all the integral curves of X have the same (maximal) interval of existence. This has two consequences. One is, of course, that we may consider only one integral curve of X, say that through e. The other, that the (maximal) existence interval of such (and all other) integral curve is $\mathbb{R}$.

To prove the latter statement, recall from the elementary theory of ODEs that 'concatenating' solutions produces solutions with larger existence intervals. Specifically, assume that X has two integral curves $\gamma_1 : (-T, T) \to M$ and $\gamma_2 : (-T, T) \to M$ (with $T > 0$) such that

$$\gamma_2(0) = \gamma_1(T^*)$$

for some $0 < T^* < T$. Then, because of uniqueness and time-translability of solutions of ODEs,[10] $\gamma_2(t) = \gamma_1(t + T^*)$ for $t \in (-T, T - T^*)$, or else

$$\gamma_1(t) = \gamma_2(t - T^*) \qquad \forall\, t \in (-T + T^*, T)\,.$$

[10] $t \mapsto \tilde\gamma_1(t) := \gamma_1(t + T^*)$ satisfies $\frac{d}{dt}\tilde\gamma_1(t) = \frac{d}{dt}\gamma_1(t + T^*) = X(\gamma_1(t + T^*)) = X(\tilde\gamma_1(t))$; hence, since $\tilde\gamma_1(0) = \gamma_2(0)$, $\tilde\gamma_1 = \gamma_2$ in the common interval of existence.

It follows that the curve $\gamma : (-T, T+T^*) \to M$ defined by

$$\begin{aligned} \gamma(t) &= \gamma_1(t) && \text{if } t \in (-T,T) \\ \gamma(t) &= \gamma_2(t-T^*) && \text{if } t \in [T, T+T^*) \end{aligned}$$

is well defined and smooth. Moreover, it is an integral curve of X with initial point $\gamma(0) = \gamma_1(0)$.[11]

The existence and uniqueness theorem ensures that there exists an integral curve γ of X through e which is defined in some interval $(-T,T)$ with $T > 0$. Concatenating it with $L_{\gamma(T/2)} \circ \gamma$ gives an integral curve through e with existence interval $(-T, 3T/2)$. Iterating this procedure, for positive and negative times, gives a sequence of intervals of existence which leave, on the left and on the right, any compact subset of $\mathbb{R}$. Thus, the maximal integral curve of X through e is defined for all times.

(ii) As noticed above, since X is left-invariant, $\Phi_t^X \circ L_g = L_g \circ \Phi_t^X$ for all g, t. Evaluating this equality at e gives $\Phi_t^X(g) = L_g \Phi_t^X(e)$.

(iii) Using the group property of flows and ii., $\Phi_{t+s}^X(e) = \Phi_t^X(\Phi_s^X(e)) = \Phi_s^X(e)\Phi_t^X(e)$; and clearly, the order of t and s may be exchanged. ■

1.2.D The Lie algebra of a Lie group. Since the set $\mathfrak{X}_L(G)$ of left-invariant vector fields on a Lie group is a Lie algebra, the linear isomorphism (1.2.5) can be used to give T_eG a Lie algebra structure as well. Specifically, define a map $[\![\ ,\ ]\!] : T_eG \times T_eG \to T_eG$ through

$$[\![\eta, \xi]\!] := [X_\eta, X_\xi](e) \qquad \forall \eta,\, \xi \,\in\, T_eG\,. \tag{1.2.6}$$

This map is bilinear (because $\xi \mapsto X_\xi$ is linear and the commutator of vector fields is bilinear) and is antisymmetric and satisfies the Jacobi identity (because $[\ ,\ ]$ has these properties). Hence, it makes T_eG into a Lie algebra.

e

Definition 1.2.7 *The Lie algebra* $(T_eG, [\![\ ,\ ]\!])$ *is called the* Lie algebra of the Lie group G*, and is denoted by* $\mathrm{lie}(G)$ *or* $\mathfrak{g}$*.*

Proposition 1.2.8 *The map* $\xi \mapsto X_\xi$ *defined by (1.2.5) is a Lie algebra isomorphism between* $(\mathfrak{g}, [\![\ ,\ ]\!])$ *and* $(\mathfrak{X}_L(G), [\ ,\ ])$*, namely*

$$X_{[\![\eta,\xi]\!]} = [X_\eta, X_\xi] \qquad \forall \eta, \xi \in \mathfrak{g}\,.$$

Proof. $X_{[\![\eta,\xi]\!]}(g) = T_eL_g \cdot [\![\eta,\xi]\!] = T_eL_g \cdot [X_\eta, X_\xi](e) = (L_g)_*[X_\eta, X_\xi](g) = [(L_g)_*X_\eta, (L_g)_*X_\xi](g) = [X_\eta, X_\xi](g)$. ■

[11] For $t < T$ this is due to the fact that γ_1 is an integral curve. For $t \geq T$: $\frac{d}{dt}\gamma(t) = \frac{d}{dt}\gamma_2(t-T^*) = X(\gamma_2(t-T^*)) = X(\gamma(t))$.

Examples:

1. Left-invariant vector fields on $G=\mathbb{R}^n$ are constant, and their commutator vanish. It follows that the Lie bracket $[\![\ ,\ ]\!]$ on $T_0\mathbb{R}^n$ is zero. In other words, the Lie algebra of $\mathbb{R}^n$ is abelian. See also Proposition 1.2.9 below.

2. As a vector space, the Lie algebra $\mathfrak{gl}(n)$ of $\mathrm{GL}(n)$ is $T_{\mathbb{I}}\mathrm{GL}(n)=\mathrm{L}(n)$ and the left-invariant vector field X_V associated to $V\in\mathrm{L}(n)$ is given by $X_V(A)=AV$ for all $A\in\mathrm{GL(n)}$. Fix $U,V\in\mathrm{L}(n)$ and recall that

$$\begin{aligned}
[X_U,X_V](A) &= -\frac{d}{dt}\Big[(\Phi_t^{X_U})_*X_V(A)\Big]_{t=0}\\
&= -\frac{d}{dt}\Big[(T\Phi_t^{X_U}\cdot X_V)(\Phi_{-t}^{X_U}(A)\Big]_{t=0}\\
&= -\frac{d}{dt}\Big[(T_{\Phi_{-t}^{X_U}(A)}\Phi_t^{X_U}\cdot X_V(\Phi_{-t}^{X_U}(A))\Big]_{t=0}.
\end{aligned}$$

Clearly, $\Phi_t^{X_U}(B)=B\exp(tU)$ for any $B\in\mathrm{GL(n)}$ and, for any $W\in T_B\mathrm{GL(n)}=\mathrm{L}(n)$, $W=\frac{d}{ds}(B+sW)|_{s=0}$. Thus

$$\begin{aligned}
T_B\Phi_t^{X_U}\cdot W &= \tfrac{d}{ds}\Phi_t^{X_U}(B+sW)|_{s=0}\\
&= \tfrac{d}{ds}(B+sW)\exp(tU)|_{s=0}\\
&= W\exp(tU)
\end{aligned}$$

and so, using also $X_V(\Phi_{-t}^{X_U}(A))=A\exp(-tU)V$,

$$\begin{aligned}
[X_U,X_V](A) &= -\frac{d}{dt}\Big[A\exp(-tU)V\exp(tU)\Big]_{t=0}.\\
&= -[A(-U)V+AVU]\\
&= A[UV-VU]\,.
\end{aligned}$$

This shows that the Lie bracket of the Lie algebra of GL(n) is the matrix commutator. Thus $\mathfrak{gl}(n)=(\mathrm{L}(n),[\ ,\ ]_-)$.

Proposition 1.2.9 *The Lie algebra of an abelian Lie group is abelian.*

Proof. We use the fact that the commutator of two vector fields is zero, $[X,Y]=0$, if and only if their flows commute, $\Phi_t^X\circ\Phi_s^Y=\Phi_s^Y\circ\Phi_t^X$ for all $t,s\in\mathbb{R}$ (Proposition A.2.7). If $\xi,\eta\in\mathfrak{g}$ then, by Proposition 1.2.6, for all $g\in G$,

$$\Phi_t^{X_\xi}\circ\Phi_s^{X_\eta}(g) \;=\; \Phi_t^{X_\xi}(g\Phi_s^{X_\eta}(e)) \;=\; g\Phi_s^{X_\eta}(e)\Phi_t^{X_\xi}(e)$$

and similarly $\Phi_s^{X_\eta}\circ\Phi_t^{X_\xi}(g) = g\Phi_t^{X_\xi}(e)\Phi_s^{X_\eta}(e)$. If the group is abelian, then these two quantities are equal and $[X_\xi,X_\eta]=0$ for all $\xi,\eta\in\mathfrak{g}$. ■

Our next goal would naturally be to determine the Lie algebras of some of the classical groups, namely, the closed Lie subgroups of GL(n). In order to do so easily, however, we may first establish a result that characterizes the Lie algebra of a Lie subgroup of a Lie group G as a certain subalgebra of the Lie algebra of G. In turn, this rests on the introduction of a further notion about vector fields, which is a weakening of the notion of push-forward and will later also play a role in the reduction under symmetry groups.

1.2.E Ψ-related vector fields. It is in general not possible to 'push-forward' a vector field X from a manifold M to another manifold N using a (smooth) non-invertible map $\Psi : M \to N$ as in the push-forward under diffeomorphisms, namely, acting on X with the tangent map $T\Psi$. If Ψ is not surjective, it is obviously impossible to define in this way the values of the 'push-forwarded' vector field in the complement of $\Psi(M) \subset N$. And if it is not injective, in which of the preimages of a point of N should $T\Psi \cdot X$ be evaluated?

$$\begin{tikzcd} M \arrow[r, "\Psi"] \arrow[d, "X"'] & N \arrow[d, squiggly, "?"] \\ TM \arrow[r, "T\Psi"] & TN \end{tikzcd}$$

However, given $X \in \mathfrak{X}(M)$, *it may happen* that there exists a vector field $Y \in \mathfrak{X}(N)$ such that

$$T\Psi \cdot X \;=\; Y \circ \Psi$$

namely

$$(T\Psi \cdot X)(m) \;=\; Y(\Psi(m)) \qquad \forall\, m \in M\,. \tag{1.2.7}$$

$$\begin{tikzcd} M \arrow[r, "\Psi"] \arrow[d, "X"'] & \Psi(M) \subseteq N \arrow[d, squiggly, "Y_{|\Psi(M)}"] \\ TM \arrow[r, "T\Psi"] & TN \end{tikzcd}$$

In such a case, X is said to be Ψ*-projectable* and Y is said to be Ψ*–related* to X (it is also common to say that X and Y are Ψ-related).[12]

Obviously, all vector fields $X \in \mathfrak{X}(M)$ are Ψ-projectable if Ψ is a diffeomorphism, and the (unique, in this case) vector field Ψ-related to X is $\Psi_* X$. If Ψ is not injective, then (1.2.7) shows that $X \in \mathfrak{X}(M)$ is Ψ-projectable if and only if $T\Psi \cdot X : M \to TN$ is constant on each fiber of Ψ. If Ψ is not surjective, and X is Ψ-projectable, then the vector field Ψ-related to X is not unique, because (1.2.7) defines it only in the points of $\Psi(M)$. In fact, this is the only reason of non-uniqueness (see Proposition 1.2.10).

Examples: 1. We want to determine the vector fields on $\mathbb{R}^2$ which are projectable under the map $\pi : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x$ (the projection onto the first factor). We may work in coordinates. Thus a vector field $X \in \mathfrak{X}(\mathbb{R}^2)$ is a map $X = (X_x, X_y) : \mathbb{R}^2 \to \mathbb{R}^2$, a vector field $Y \in \mathfrak{X}(\mathbb{R})$ is a function $Y : \mathbb{R} \to \mathbb{R}$ and the tangent map $T_{(x,y)}\pi$ is the Jacobian matrix $\pi'(x, y) = (1, 0)$. Thus,

$$\pi'(x, y)X(x, y) = (1, 0)\begin{pmatrix} X_x(x, y) \\ X_y(x, y) \end{pmatrix} = X_x(x, y)\,.$$

It follows that X is π-projectable if and only if its x-component is independent of y. In such a case, the vector field Y on $\mathbb{R}$ which is π-related to X is

$$Y(x) = X_x(x)\,.$$

[12] The term 'projectable' appears in the literature, but it is not of widespread use. There are understandable reasons for this. On the one hand, it is convenient to have a term that says that a vector field has the property that there exist vector field(s) Ψ-related to it. However, 'Ψ-projectable' suggests that Ψ ia a projection, namely a surjection, or even more a surjective submersion, as in Example 1 below. But it does not convey the right intuition if Ψ is an immersion, as in Example 2 below; perhaps 'injectable' would work better in this case. A name which is good for all cases seems to be missing.

Note that there are no conditions on the y-component of X: if $\Psi : M \to N$ is not a diffeomorphism, then a vector field on N may be Ψ-related to different vector fields on M.

2. Consider the immersion $\iota : \mathbb{R} \to \mathbb{R}^2$, $t \mapsto (\cos t, \sin t)$. ι is neither surjective (its image is the unit circle) nor injective ($\iota^{-1}(\cos t, \sin t) = \{t + 2k\pi : k \in \mathbb{Z}\}$). We work in coordinates, so that $X \in \mathfrak{X}(\mathbb{R})$ is a function $\mathbb{R} \to \mathbb{R}$ etc. Since

$$\iota'(t)X(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} X(t) = \begin{pmatrix} -X(t)\sin t \\ X(t)\cos t \end{pmatrix}$$

X is ι-projectable if and only if $(-X(t)\sin t, X(t)\cos t)$ is constant on the fibers $\{t + 2k\pi : k \in \mathbb{Z}\}$ of ι, namely, if and only if $X : \mathbb{R} \to \mathbb{R}$ is 2π-periodic. In such a case, if $\in \mathfrak{X}(\mathbb{R}^2)$ is ι-related to X, then

$$Y(\cos t, \sin t) = (-X(t)\sin t, X(t)\cos(t)) \qquad \forall\, t \in \mathbb{R}\,.$$

For instance, if $X(t) = \cos t$, then a vector field Y on $\mathbb{R}^2$ which is ι-related to X is

$$Y(x, y) = (-xy, x^2)\,.$$

(But if Z is any vector field in $\mathbb{R}^2$ that vanishes on the unit circle, then $Y + Z$ is ι-related to X as well).

We now see some properties of Ψ-related vector fields that we will need later.

Proposition 1.2.10 *Let $\Psi : M \to N$ be a (smooth) map between two manifolds M and N and let $X \in \mathfrak{X}(M)$ be a Ψ-projectable vector field.*

i. If Ψ is surjective, then the vector field Ψ-related to X is unique.

ii. Let $Y \in \mathfrak{X}(N)$. Then Y is Ψ-related to X if and only if

$$\Phi_t^Y \circ \Psi = \Psi \circ \Phi_t^X \qquad \forall t\,.$$

iii. Consider now another Ψ-projectable vector field $X' \in \mathfrak{X}(M)$. Let $Y \in \mathfrak{X}(N)$ be Ψ-related to X and $Y' \in \mathfrak{X}(N)$ be Ψ-related to X'. Then, $[X, X']$ is Ψ-projectable and $[Y, Y']$ is Ψ-related to it.

Proof. (i) If Ψ is surjective then for each $n \in N$ there is $m \in M$ such that $n = \Psi(m)$ and, by (1.2.7), $Y(n)$ equals $(T\Psi \cdot X)(m)$.

(ii) Suppose that $\Phi_t^Y \circ \Psi = \Psi \circ \Phi_t^X \quad \forall t$. Evaluating at $m \in M$ and taking $\frac{d}{dt}\big|_{t=0}$ yields $Y(\Psi(m)) = T_m\Psi X(m)$ so (1.2.7) holds and X and Y are Ψ-related.

Now assume that X and Y are Ψ-related. The condition $\Phi_t^Y \circ \Psi = \Psi \circ \Phi_t^X \quad \forall t$ is equivalent to saying that if $t \mapsto m_t$ is an integral curve of X, then $t \mapsto n_t := \Psi(m_t)$ is an integral curve of Y. In fact, $\dot n_t = \frac{d}{dt}\Psi(m_t) = T_{m_t}\Psi \cdot \dot m_t = T_{m_t}\Psi \cdot X(m_t) = (T\Psi \cdot X)(m_t) = Y(\Psi(m_t))$.

(iii) We need the following

Lemma 1.2.11 *Let $\Psi : M \to N$. $Y \in \mathfrak{X}(N)$ is Ψ-related to $X \in \mathfrak{X}(M)$ if and only if*

$$L_X(g \circ \Psi) = (L_Y g) \circ \Psi \qquad \forall\, g \in C^\infty(N)\,. \tag{1.2.8}$$

Proof of the Lemma. For any $m \in M$,

$$\begin{aligned} L_X(g\circ\Psi)(m) &= T_m(g\circ\Psi)\cdot X(m) = T_{\Psi(m)}g \cdot T_m\Psi\cdot X(m) \\ (L_Y g)(\Psi(m)) &= T_{\Psi(m)}g\cdot Y(\Psi(m))\,. \end{aligned}$$

If Y is Ψ-related to X then the right hand sides of these expressions are equal and this proves (1.2.8). Conversely, if (1.2.8) is satisfied, then the previous equalities imply $T_{\Psi(m)}g\cdot T_m\Psi\cdot X(m) = T_{\Psi(m)}g\cdot Y(\Psi(m))$ for all $g \in C^\infty(M)$ and all $m \in M$. By the linearity of the tangent map, and the definition (A.1.2) of the exterior derivative of a real function, this implies

$$\langle dg(\Psi(m)), T_m\Psi\cdot X(m) - Y(\Psi(m))\rangle_{\Psi(m)} = 0 \quad \forall\, m\in M\,,\ g \in C^\infty(N)\,.$$

By the arbitrariness of g, this implies that, for each $m \in M$, the vector $T_m\Psi\cdot X(m) - Y(\Psi(m)) \in T_\Psi(m)N$ is zero. ■

Instead of proving that $T\Psi\cdot[X,X'] = [Y,Y']\circ\Psi$ we may thus prove that $L_{[X,X']}(g\circ\Psi) = (L_{[Y,Y']}g)\circ\Psi$ for all $g \in C^\infty(N)$. This is verified with an algebraic-like computation analogous to that used to prove the naturalness of the Lie brackets under push-forward in Proposition A.2.6. ■

Exercises 1.2.3 *(i) Let $\pi : \mathbb{R} \to \mathbb{T}^1$, $x \mapsto x\mathrm{mod}1$. Which vector fields on $\mathbb{R}$ are π-projectable to vector fields on $\mathbb{T}^1$?*

1.2.F Lie algebras of subgroups. We use the following notation: if $G, H, \dots$ are Lie groups, then $\mathfrak{g}, \mathfrak{h}, \dots$ are their Lie algebras with Lie brackets $[\![\ ,\ ]\!]^{\mathfrak{g}}, [\![\ ,\ ]\!]^{\mathfrak{h}}, \dots$ and, when it is necessary to avoid ambiguities, $e_G, e_H, \dots$ are their identity elements. Also, $X^G_\xi, X^H_\xi, \dots$ denote the left-invariant vector fields on $G, H, \dots$ associated to an element ξ of their Lie algebra.

Lemma 1.2.12 *Assume that $\Psi : G \to H$ is a Lie group homomorphism. Then, for any $\xi \in \mathfrak{g}$, X^G_ξ and $X^H_{T_{e_G}\Psi\cdot\xi}$ are Ψ-related.*

Proof. Since Ψ is a group homomorphism, $\Psi(L_g g') = \Psi(gg') = L_{\Psi(g)}\Psi(g')$ $\forall g, g' \in G$, or

$$\Psi\circ L_g \; = \; L_{\Psi(g)}\circ\Psi \qquad \forall g \in G\,.$$

Computing the tangent map at e_G of both quantities, and using the chain rule and $e_H = \Psi(e_G)$, gives

$$T_g\Psi\cdot T_{e_G}L_g = T_{e_H}L_{\Psi(g)}\cdot T_{e_G}\Psi\,.$$

Now, for any $\xi \in \mathfrak{g}$,

$$\begin{aligned} T_g\Psi \cdot T_{e_G}L_g \cdot \xi &= T_g\Psi \cdot X^G_\xi(g) \\ T_{e_H}L_{\Psi(g)} \cdot T_{e_G}\Psi \cdot \xi &= X^H_{T_{e_G}\Psi\cdot\xi}(\Psi(g))\,. \end{aligned}$$

It follows that

$$T_g\Psi \cdot X^G_\xi(g) = X^H_{T_{e_G}\Psi\cdot\xi}(\Psi(g)) \quad \forall g \in G$$

which proves the claim. ■

Proposition 1.2.13 *If $\Psi : G \to H$ is a homomorphism (isomorphism) of Lie groups, then $T_{e_G}\Psi : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism (isomorphism) of Lie algebras.*

Proof. $T_{e_G}\Psi$ is a linear map between $T_{e_G}G$ and $T_{e_H}H$ and we only need to check that

$$T_{e_G}\Psi \cdot [\![\eta,\xi]\!]^{\mathfrak{g}} \;=\; [\![T_{e_G}\Psi\cdot\eta, T_{e_G}\Psi\cdot\xi]\!]^{\mathfrak{h}} \qquad \forall\,\eta,\xi \in T_{e_G}G\,.$$

Using the definition of the brackets of $\mathfrak{g}$ and $\mathfrak{h}$, Lemma 1.2.11, Proposition 1.2.10, and $\Psi(e_G) = e_H$, we compute

$$\begin{aligned} T_{e_G}\Psi \cdot [\![\eta,\xi]\!]^{\mathfrak{g}} &= T_{e_G}\Psi \cdot [X^G_\eta, X^G_\xi](e_G) \\ &= [X^H_{T_e\Psi\cdot\eta}, X^H_{T_e\Psi\cdot\xi}](e_H) \\ &= [\![T_{e_G}\Psi\cdot\eta, T_{e_G}\Psi\cdot\xi]\!]^{\mathfrak{h}}\,. \end{aligned}$$

Lastly, the statement concerning isomorphisms is obvious: if Ψ is a diffeomorphism, then its tangent map is a linear isomorphism. ■

> **Remark:** As the proof makes clear, this means that, if $\Psi : G \to H$ is a Lie group homomorphism, then $T_{e_G}\Psi(T_{e_G}G)$, which is a linear subspace of $T_{e_H}H$, is a subalgebra of $\mathfrak{h}$: namely, the restriction to $T_{e_G}\Psi(T_{e_G}G) \times T_{e_G}\Psi(T_{e_G}G)$ of the Lie product $[\![\ ,\ ]\!]^{\mathfrak{h}}$ on $T_{e_H}H$ defines a Lie product on $T_{e_G}\Psi(T_eG)$.

Recall from section 1.1.E that a Lie subgroup H of a Lie group G is a Lie group which is injectively immersed into G, the immersion being a Lie group homomorphism. There is no request, in this definition, about the relationship between the Lie algebra of H (as a Lie group in itself) and that of G. However, Proposition 1.2.13 implies that:

Proposition 1.2.14 *Let G be a Lie group with Lie algebra $\mathfrak{g} = \big(T_eG, [\![\ ,\ ]\!]^{\mathfrak{g}}\big)$ and let S be a Lie subgroup of G. Define*

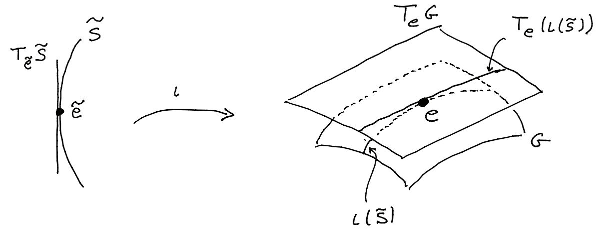
$$\mathfrak{g}|_{T_e(S)} := \big(T_eS, [\![\ ,\ ]\!]^{\mathfrak{g}}\big|_{T_eS\times T_eS}\big)\,.$$

Then:

i. $\mathfrak{g}|_{T_e(S)}$ *is a Lie subalgebra of* $\mathfrak{g}$.

ii. The Lie algebra of S is isomorphic, as a Lie algebra, to $\mathfrak{g}|_{T_e(S)}$

Proof. It may be clearer to regard, as in section 1.1.E, S as given by an injective immersion $\iota : \tilde{S} \to G$, which is also a Lie group homomorphism, of a Lie group $\tilde{S}$ into G. Thus $\iota(\tilde{S}) = S$ and $\iota : \tilde{S} \to \iota(\tilde{S}) \subset G$ is a diffeomorphism, and hence a Lie group isomorphism. Let $\tilde{e}$ be the identity element of $\tilde{S}$ and $\tilde{\mathfrak{s}}$ the Lie algebra of $\tilde{S}$.



By Proposition 1.2.13 and the Remark following it, $T_{\tilde{e}}\iota(T_{\tilde{e}}\tilde{S})$ is a Lie subalgebra of $\mathfrak{g}$ and

$$T_{\tilde{e}}\iota \,:\, \tilde{\mathfrak{s}} \to \Big(T_{\tilde{e}}\iota(T_{\tilde{e}}\tilde{S})\,,\, [\![\,,\,]\!]^{\mathfrak{g}}\big|_{T_{\tilde{e}}\iota(T_{\tilde{e}}\tilde{S})\times T_{\tilde{e}}\iota(T_{\tilde{e}}\tilde{S})}\Big)$$

is a Lie algebra isomorphism. The proof is concluded by observing that, since ι is an immersion, $T_{\tilde{e}}\iota(T_{\tilde{e}}\tilde{S}) = T_{e_G}(\iota(\tilde{S}))$. ∎

Thus, in particular, the Lie algebra of a Lie subgroup S of $\mathrm{GL}(n)$ is the subspace $T_{\mathbb{I}}S$ of $T_{\mathbb{I}}\mathrm{GL(n)} = L(n)$, with the matrix commutator as product.

Examples:

1. $\mathfrak{o}(n)$. A matrix $V \in \mathrm{L}(n)$ belongs to $T_{\mathbb{I}}\mathrm{O}(n)$ if and only if there exists a curve $t \mapsto A_t \in \mathrm{O}(n)$ such that $A_0 = \mathbb{I}$ e $V = \dot{A}_0$. The matrix $t \mapsto A_t \in \mathrm{O}(n)$ if and only if $A_tA_t^T = \mathbb{I}$. Therefore, $t \mapsto A_t \in \mathrm{O}(n)$ for all t if and only if $A_0 = \mathbb{I}$ and $0 = \frac{d}{dt}(A_tA_t^T) \;=\; A_t\dot{A}_t^T + \dot{A}_tA_t^T$ for all t. If $A_0 = \mathbb{I}$, this implies $\dot{A}_0 + \dot{A}_0^T = 0$. Thus, any matrix $V \in T_{\mathbb{I}}\mathrm{O}(n)$ is antisymmetric, and $T_{\mathbb{I}}\mathrm{O}(n)$ is a vector subspace of skew (n). But skew (n) and $\mathrm{O}(n)$ have the same dimension and we conclude that $T_{\mathbb{I}}\mathrm{O}(n) = \mathrm{skew}\,(n)$ and

$$\mathfrak{o}(n) \;=\; (\mathrm{skew}\,(n)\,,\, [\,,\,]_-)\,.$$

In particular

$$\mathfrak{so}(3) \;=\; (\mathrm{skew}\,3\,,\, [\,,\,]_-)\,.$$

On account of the isomorphism $\hat{}: \mathbb{R}^3 \to \mathrm{skew}\,(3)$, $\mathfrak{so}(3)$ is also isomorphic, as a Lie algebra, to $(\mathbb{R}^3\,,\,\times)$.

2. $\mathfrak{sl}(n)$. $\mathrm{SL}(n)$ is the subgroup of $\mathrm{GL(n)}$ formed by the matrices with determinant 1, and has dimension $n^2 - 1$. A matrix $V \in \mathrm{L}(n)$ belongs to $T_{\mathbb{I}}\mathrm{SL}(n)$ if and only if it equals $\dot{A}_0$ for a curve $t \mapsto A_t \in \mathrm{SL}(n)$, namely a curve in $\mathrm{L}(n)$ such

that $A_0 = \mathbb{I}$ and $\frac{d}{dt}\det(A_t) = 0$ for all t. For small $|t|$, $A_t = \mathbb{I} + t\dot{A}_0 + \mathcal{O}(t^2)$. Thus

$$\det A_t = \det(\mathbb{I} + t\dot{A}_0 + \mathcal{O}(t^2)) = 1 + t\,\mathrm{tr}\,(\dot{A}_0) + \mathcal{O}(t^2)$$

(see the example of SU(n) in section 1.1.B) and so $\frac{d}{dt}A_t|_{t=0} = \mathrm{tr}\,(\dot{A}_0)$. This shows that $T_{\mathbb{I}}\mathrm{SL}(n) \subseteq \{V \in \mathrm{L}(n) : \mathrm{tr}\,(V) = 0\}$ and, since the dimensions are the same,

$$\mathfrak{sl}(n) = (\text{Traceless n} \times \text{n matrices}\,, [\ ,\ ]_-)\,.$$

3. $\mathfrak{s}^3$. As in sections 1.1.E and 1.1.G, let us regard the group S^3 of unit quaternions, and its tangent spaces, as submanifolds of $\mathbb{R}^4$. The identity element is $(1,\overline{0})$ and $T_{(1,\overline{0})}S^3 = \{(0,\overline{v}) : \overline{v} \in \mathbb{R}^3\}$. Since the Euler parameter map $\mathcal{E} : S^3 \to \mathrm{SO}(3)$ (see (1.1.10)) is a Lie group homomorphism its tangent map $T_{e_{S^3}}\mathcal{E} : \mathfrak{s}^3 \to \mathrm{skew}\,(3)$ is a Lie algebra homomorphism—and in fact an isomorphism because $T_{e_{S^3}}$ and skew (3) are both 3-dimensional. This isomorphism was actually computed as part of the proof of Proposition 1.1.13, and is given by

$$T_{(1,\overline{0})}\mathcal{E} \cdot (0,\overline{v}) = 2\hat{v}\,.$$

Thus, the Lie bracket of $\mathfrak{s}^3$ satisfies $T_{(1,\overline{0})}\mathcal{E} \cdot [\![(0,\overline{u}),(0,\overline{v})]\!]^{\mathfrak{s}^3} = 2[T_{(1,\overline{0})}\mathcal{E} \cdot (0,\overline{u}), T_{(1,\overline{0})}\mathcal{E} \cdot (0,\overline{v})]_- = 8\widehat{\overline{u} \times \overline{v}}$ and we conclude that $[\![(0,\overline{u}),(0,\overline{v})]\!]^{\mathfrak{s}^3} = (0, 4\overline{u} \times \overline{v})$.

Exercises 1.2.4 *(i) Show that* $\mathfrak{u}(n) = \{U \in \mathrm{L}(n,\mathbb{C}) : U + U^* = 0\}$.

(ii) Here is another proof of the fact that the Lie algebra of an abelian Lie group is abelian (Proposition 1.2.9). Consider the inversion $I : G \to G$, $g \mapsto g^{-1}$. *Show that*

(a) $T_eI \cdot [\![\eta,\xi]\!] = -[\![\xi,\eta]\!]$ *for all* $\eta,\xi \in \mathfrak{g}$.
(b) If G is abelian, I is a Lie group isomorphism.

Deduce that, if G is abelian, then $[\![\eta,\xi]\!] = -[\![\eta,\xi]\!]$ *for all* $\eta,\xi \in \mathfrak{g}$.

1.3 The exponential map

1.3.A Definition, examples and first properties. Recall that, if $\xi \in \mathfrak{g}$, then the left-invariant vector field X_ξ is given by $X_\xi(g) = T_eL_g \cdot \xi$ and its flow, which is defined for all $t \in \mathbb{R}$, satisfies $\Phi_t^{X_\xi}(g) = g\Phi_t^X(e)$ for all $g \in G$. For each $t \in \mathbb{R}$, $\Phi_t^{X_\xi}$ is a diffeomorphism of G onto itself.

Definition 1.3.15 *The exponential map of a Lie group G is the map*

$$\exp_G : \mathfrak{g} \to G\,, \qquad \xi \mapsto \exp_G(\xi) := \Phi_1^{X_\xi}(e)\,.$$

If G is fixed, we will routinely write exp instead of $\exp_G$. Proposition 1.2.6 implies that

$$\exp((t+s)\xi) = \exp(t\xi)\exp(s\xi) = \exp(s\xi)\exp(t\xi) \qquad \forall\, t,s \in \mathbb{R}\,,\ \xi \in \mathfrak{g}$$

but, unless G is abelian, in general $\exp(\xi + \eta) \neq \exp(\xi)\exp(\eta)$ for $\xi,\eta \in \mathfrak{g}$.

Examples: 1. $\mathbb{R}^n$. Since $T_0\mathbb{R}^n = \mathbb{R}^n$, the exponential map $\exp_{\mathbb{R}^n} : T_0\mathbb{R}^n \to \mathbb{R}^n$ of $\mathbb{R}^n$ can be regarded as a map of $\mathbb{R}^n$ into itself. The left-invariant vector fields of $\mathbb{R}^n$ are constant, $X_\xi(x) = \xi$ for all $\xi \in \mathbb{R}^n$ (the Lie algebra) and $x \in \mathbb{R}^n$ (the group). Hence $\Phi_t^{X_\xi}(x) = x + t\xi$ and

$$\exp_{\mathbb{R}^n}(\xi) \;=\; \xi \qquad \forall \xi \in \mathbb{R}^n$$

namely

$$\exp_{\mathbb{R}^n} = \mathrm{id}_{\mathbb{R}^n}\,.$$

2. $\mathrm{GL}(n)$. For any $V \in T_{\mathbb{I}}\mathrm{GL(n)} = \mathrm{L}(n)$, $X_V(A) = AV$ for all $A \in \mathrm{GL(n)}$. As we know (see Example 2. in section 1.2.D) $\Phi_t^{X_V}(A) = A\exp(Vt)$, where exp denotes the matrix exponential. Hence,

$$\exp_{\mathrm{GL(n)}} : \mathfrak{gl}(n) = \mathrm{L}(n) \to \mathrm{GL(n)}$$

is the matrix exponential.

Proposition 1.3.16 *The exponential map* $\exp_G : \mathfrak{g} \to G$ *of a Lie group* G*:*

i. Is a smooth map.

ii. Is a local diffeomorphism at $0 \in \mathfrak{g}$.

Proof. (i) If a vector field depends smoothly on parameters, then the maps at fixed t of its flow are smooth maps of the parameters as well (see the Exercises below). X_ξ is a linear (hence smooth) function of ξ, and $\xi \mapsto \Phi_1^{X_\xi}(e)$ is thus smooth.

(ii) Since $\exp(0) = e$, by the inverse function theorem this is equivalent to the fact that $T_0 \exp : T_0\mathfrak{g} \equiv \mathfrak{g} \to T_eG = \mathfrak{g}$ is invertible. Since $\Phi_t^{X_\xi} = \Phi_1^{X_{t\xi}}$ for all t and ξ (see the Exercises) we compute

$$T_0 \exp \cdot \xi \;=\; \frac{d}{dt}\exp(t\xi)_{|t=0} \;=\; \frac{d}{dt}\Phi_1^{t\xi}(0)_{|t=0} \;=\; \frac{d}{dt}\Phi_t^{\xi}(0)_{|t=0} \;=\; X_\xi(0) \;=\; \xi$$

and so $T_0 \exp = \mathrm{id}_{\mathfrak{g}}$. ■

Thus, the exponential map of a Lie group G is a diffeomorphism between a neighbourhood of $0 \in \mathfrak{g}$ and a neighbourhood of $e \in G$. As such, it allows to model the local structure of the Lie group near its identity element by means of that of the algebra. However, $\exp : \mathfrak{g} \to G$:

- may be not surjective
- may be not injective
- may be not a local diffeomorphism at a point $\xi \neq 0$.

Let us say something about the (lack of) surjectivity of exp. Since it is a continuous map and $\mathfrak{g}$ is connected, $\exp(\mathfrak{g})$ is connected. Hence, the exponential map of a nonconnected group is never surjective. If G is not connected, then, given that $\exp(0) = e$, the image $\exp(\mathfrak{g})$ of the exponential map is contained in the connected component G_0 of G that contains the identity. However, exp need not be surjective even on G_0.

Examples: 1. We will see later that the exponential map of the classical Lie groups is the (restriction of) the matrix exponential. Thus, Proposition 1.1.11 means that the exponential map of SO(3) is surjective.

2. Consider the connected component of GL(n) that contains $\mathbb{I}$, namely $\mathrm{GL}_+(n) = \{A \in \mathrm{L}(n) : \det A > 0\}$. Its Lie algebra and exponential map obviously coincide with those of GL(n). Thus, the exponential map is the matrix exponential $\exp : \mathrm{L}(n) \to \mathrm{GL}_+(n)$ and its surjectivity amounts to the fact that for any $A \in \mathrm{GL}_+(n,\mathbb{R})$ there exists a matrix $L \in \mathrm{L}(n,\mathbb{R})$, called a *real logarithm* of A, such that $A = \exp L$. It is known from linear algebra that a matrix $A \in \mathrm{GL(n)}$ has a (possibly non unique) real logarithm if and only if in its Jordan normal form each block relative to a negative real eigenvalue appears an even number of times.[13] For instance, the matrix $\mathrm{diag}\,(-1,-2)$ belongs to $\mathrm{GL}_+(2)$ and has no real logarithm. Thus, for any $n \geq 2$, $\exp_{\mathrm{GL(n)}}$ is not surjective.

We quote without proof the following result:

Proposition 1.3.17 *The exponential map of a Lie group G is surjective in the following cases:*

i. G is compact and connected.

ii. G is abelian and connected.

However, these are only sufficient conditions.

Exercises 1.3.5 *(i) Which is the exponential map of $\mathbb{R}_*$ (non-zero real numbers, with the real multiplication)?*

(ii) The smooth dependence on parameters of the solutions of ODEs is a local question and can be investigated in coordinates. Consider a smooth map $\mathbb{R}^n \times \mathbb{R}^m \ni (z,c) \mapsto X(z,c) = X_c(z)$. Assume that, for c in a neighbourhood of $\overline{c} \in \mathbb{R}^m$, all vector fields X_c are complete. Show that, for each z_0 and t, the map $\mathbb{R}^m \ni c \mapsto \Phi_t^{X_c}(z_0) \in \mathbb{R}^n$ is smooth at $\overline{c}$. [Hints: consider the flow of the ODE $\dot z = X(z,c)$, $\dot c = 0$ in $\mathbb{R}^{n+m} \ni (z,c)$].

(iii) Show that, if X is a (complete) vector field and $k \in \mathbb{R}$, then $\Phi_{kt}^X = \Phi_t^{kX}$ for all t. [Hints: Using the chain rule, verify that $t \mapsto m_t := \Phi_{kt}^X(m_0)$ is an integral curve of kX; observe that it has initial datum m_0 at $t = 0$ and invoke the uniqueness.]

(iv) Verify that $\exp(t\xi) = \Phi_t^{X_\xi}(e)$ for all $\xi \in \mathfrak{g}$ and $t \in \mathbb{R}$.

(v) Find a real logarithm of the matrix $\mathrm{diag}\,(-1,-1)$. Is it unique? [Hint: compute $\exp\begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix}$ with $\beta \in \mathbb{R}$].

(vi) Show that if $[\![\xi,\eta]\!] = 0$ then $\exp(\xi+\eta) = \exp(\xi)\exp(\eta)$.

[13] More exhaustively: Let $B \in \mathrm{L}(n,\mathbb{C})$. Then $B = \exp(L)$ with a complex matrix L if and only if B is invertible. And $B = \exp(L)$ with a real matrix L if and only if B is invertible and, for each negative eigenvalue λ of B, each Jordan block relative to λ is repeated an even number of times.

1.3.B The exponential map and Lie group homomorphisms.

Proposition 1.3.18 *Let G, H be Lie groups and $\Psi : G \to H$ a Lie group homomorphism. Then*

$$\Psi \circ \exp_G \;=\; \exp_H \circ T_{e_G}\Psi\,.$$

$$\begin{tikzcd} G \arrow[r, "\Psi"] & H \\ \mathfrak{g} \arrow[u, "\exp_G"] \arrow[r, "T_{e_G}\Psi"'] & \mathfrak{h} \arrow[u, "\exp_H"'] \end{tikzcd}$$

Proof. Fix $\xi \in \mathfrak{g}$. Since Ψ is a Lie group homomorphism X^G_ξ and $X^H_{T_{e_G}\Psi\cdot\xi}$ are Ψ-related (Lemma 1.2.12) and hence $\Phi_t^{X^H_{T_e\Psi\cdot\xi}}\circ\Psi = \Psi\circ\Phi_t^{X^G_\xi}$ for all t (Proposition 1.2.10, statement ii). It follows that $\exp_H(T_e\Psi\cdot\xi) = \Phi_1^{X^H_{T_e\Psi\cdot\xi}}(e_H) = \Phi_1^{X^H_{T_e\Psi\cdot\xi}}\circ\Psi(e_G)$ and $\Psi\circ\exp_G(\xi) = \Psi\circ\Phi_1^{X^G_\xi}(e_G)$ are equal. ■

Proposition 1.3.18 has important consequences. First, it implies that the exponential map of a Lie subgroup S of a Lie group G is the restriction to T_eS of the exponential map of G. Precisely:

Corollary 1.3.19 *Let $\iota : \tilde{S} \to G$ be a Lie subgroup of G. Then, $\iota\circ\exp_{\tilde{S}} = \exp_G \circ T_{\tilde{e}}\iota$.*

Proof. The immersion ι is a Lie group homomorphism. ■

Thus, the *exponential map of all Lie subgroups of* $\mathrm{GL}(n)$ *is the matrix exponential.* We see now a few other examples about the exponential maps.

Examples: 1. $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$. Since $\mathbb{T}^1$ is 1-dimensional, its Lie algebra is $(\mathbb{R}, +)$.[14] The covering map

$$\pi : \mathbb{R} \to \mathbb{T}^1\,, \qquad x \mapsto x \bmod 1\,,$$

is a Lie group homomorphism (immediate to verify). Hence,

$$\exp_{\mathbb{T}^1} \circ T_0\pi = \pi\circ\exp_{\mathbb{R}}\;.$$

But $T_0\pi : T_o\mathbb{R} \sim \mathbb{R} \to T_e\mathbb{T}^1 \sim \mathbb{R}$ is the identity (in suitable coordinates near $0 \in \mathbb{R}$ and $e = 0 \bmod 1 \in \mathbb{T}^1$ the projection π is the identity). Thus

$$\exp_{\mathbb{T}^1} = \pi\circ\exp_{\mathbb{R}} = \pi\circ\mathrm{id} = \pi\,.$$

In particular, $\exp_{\mathbb{T}^1}$ is surjective but not injective.

2. Consider the direct product of two Lie groups G_1 and G_2, $G = G_1 \times G_2$ with product is $(g_1, g_2)(h_1, h_2) = (g_1g_2, h_1h_2)$. Everything factorizes: the tangent spaces $T_{g_1,g_2}(G_1\times G_2)$ are isomorphic to the direct sum $T_{g_1}G_1 \oplus T_{g_2}G_2$ and vector

[14]This follows from the fact that $\mathbb{T}^1$ is abelian, but it is trivial in this case because every 1-dimensional Lie algebra is abelian.

fields on G can be identified with pairs (X_1, X_2) of vector fields $X_1 \in \mathcal{X}(G_1)$, $X_2 \in \mathcal{X}(G_2)$. Since left-translations in G factorize as well, (X_1, X_2) is left-invariant if and only if X_1 and X_2 are. Thus, the Lie algebra $\mathrm{lie}(G_1 \times G_2)$ splits as the direct sum of $\mathrm{lie}(G_1)$ and $\mathrm{lie}(G_2)$ and the exponential map factorizes as well:

$$\exp_{G_1 \times G_2}(\xi_1, \xi_2) = (\exp_{G_1}(\xi_1), \exp_{G_2}(\xi_2)) \qquad \forall \xi_1 \in \mathrm{lie}G_1,\ \xi_2 \in \mathrm{lie}G_2\,.$$

3. $\mathbb{T}^n = \mathbb{T}^1 \times \cdots \times \mathbb{T}^1$ has (abelian) Lie algebra $\mathbb{R}^n$ and exponential map $\exp_{\mathbb{T}^n}(\xi) = \xi \,\mathrm{mod} 1$. Again, surjective but not injective.

Remark:

Exercises 1.3.6 *(i) Show that, if* exp *denotes the matrix exponential, then*

$$\det(\exp V) = e^{\mathrm{tr}\, V} \qquad \forall\, V \in \mathrm{L}(n)\,.$$

[Hints: $\det : \mathrm{GL(n)} \to \mathbb{R}_*$ *is a Lie group homomorphism and the exponential maps of* GL(n) *and* $\mathbb{R}_*$ *are .....]*

1.3.C One-parameter subgroups.

Definition 1.3.20 *A* one-parameter subgroup *of a Lie group G is a Lie group homomorphism $\gamma : \mathbb{R} \to G$.*

Thus, a one-parameter subgroup of G is a (smooth) curve $\gamma : \mathbb{R} \to G$ through the identity ($\gamma(0) = e$) and such that

$$\gamma(t + s) = \gamma(t)\gamma(s) \qquad \forall t, s \in \mathbb{R}\,.$$

By Proposition 1.2.6, for any $\xi \in \mathfrak{g}$, the integral curve $t \mapsto \Phi_t^{X_\xi}(e)$ of X_ξ through the identity e satisfies $\Phi_{t+s}^{X_\xi}(e) = \Phi_t^{X_\xi}(e)\Phi_s^{X_\xi}(e)$ $\forall t, s$ and is a one-parameter subgroup. Equivalently, since $\Phi_t^{X_\xi}(e) = \exp(t\xi)$ (see Exercise 1.3.5.iv), for any $\xi \in \mathfrak{g}$ the curve $t \mapsto \exp(t\xi)$ is a one-parameter subgroup. In fact, all one-parameter subgroups are of this type:

Proposition 1.3.21 *If $\gamma : \mathbb{R} \to G$ is a one-parameter subgroup, then*

$$\gamma(t) = \exp(t\xi) \qquad \forall t \in \mathbb{R}$$

with a unique $\xi \in \mathfrak{g}$.

Proof. Since γ is a group homomorphism, $\gamma(0) = e$. Hence

$$\xi := \gamma'(0) \in T_e G\,.$$

Since $\gamma(0) = e$, by the definition of flow of a vector field, $\gamma(t) = \Phi_t^{X_\xi}(e)$ is equivalent to $\gamma'(t) = X_\xi(\gamma(t))$ for all t. And in fact

$$\begin{aligned}\gamma'(t) &= \frac{d}{ds}\gamma(t+s)|_{s=0} = \frac{d}{ds}\big(\gamma(t)\gamma(s)\big)|_{s=0} = \frac{d}{ds}L_{\gamma(t)}\gamma(s)\big)|_{s=0} \\ &= T_{\gamma(s)}L_{\gamma(t)} \cdot \gamma'(s)|_{s=0} = T_e L\gamma(t) \cdot \xi = X_\xi(\gamma(t))\,.\end{aligned}$$

If $\xi, \eta \in \mathfrak{g}$ are such that $\exp(t\xi) = \exp(t\eta)$ for all t, then $\xi = \frac{d}{dt}\exp(t\xi)|_{t=0} = \frac{d}{dt}\exp(t\eta)|_{t=0} = \eta$. ■

Definition 1.3.22 *The vector $\xi \in \mathfrak{g}$ is the* (infinitesimal) generator *of the one-parameter subgroup $t \mapsto \exp(t\xi)$.*

One-parameter subgroups are an important tool in the study of Lie group and of dynamical systems invariant under Lie group actions.

Exercises 1.3.7 *(i) Show that the image of the one-parameter subgroup of $\mathbb{T}^2$ generated by a vector $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$ is a circle if ξ_1 and ξ_2 are linearly dependent over $\mathbb{Q}$ and is dense in $\mathbb{T}^2$ otherwise.*

(ii) Show on with an example that two one-parameter subgroups with different generators may intersect at (isolated) points other than e. [Hint: The previous exercise]

1.3.D On the relation between Lie algebras and Lie groups. We give now a quick look at the correspondence between Lie groups and Lie algebras: every Lie group determines its Lie algebra, but to which extent does the algebra determine the group? Obviously, since (as we have already remarked) the image of the exponential map is connected, this question can be asked only for *connected* groups. Furthermore, it should be put between isomorphism classes of Lie algebras and isomorphism classes of Lie groups.

Proposition 1.2.13 implies that isomorphic Lie groups have isomorphic Lie algebras. However, simple examples show that there exist non-isomorphic Lie groups which do have isomorphic Lie algebras. For instance, $\mathbb{R}^n$ and $\mathbb{T}^n$ are not isomorphic Lie groups but have the same Lie algebra $\mathbb{R}^n$. One may think that the problem here is that $\mathbb{R}^n$ is not compact and $\mathbb{T}^n$ is, but S^3 and $\mathrm{SO}(3)$, which are both compact, are not isomorphic as Lie groups and nevertheless have isomorphic Lie algebras ($\mathrm{lie}(S^3) = (\mathbb{R}^3, 4\times)$), see example 3. of section 1.2.A, and $(\mathrm{skew}\,(3), [\ ,\ ]_-)$). The reason is a different one: simply connectedness. $\mathbb{R}^n$ and S^3 are simply connected but $\mathbb{T}^n$ and $\mathrm{SO}(3)$ are not.

Proposition 1.3.23 *There is a one-to-one correspondance between isomorphism classes of (connected*[15] *and) simply connected Lie groups and isomorphism classes of Lie algebras.*

[15]Usually, path-connectedness and hence connectedness is required in the definition of simply-connectedness, so specifying ‘connected’ here might not be necessary.

More generally:

Proposition 1.3.24 *Given a Lie Algebra $\mathcal{A}$, the connected Lie groups whose Lie algebra is isomorphic to $\mathcal{A}$ are, up to isomorphisms: a simply connected Lie group G whose Lie algebra is $\mathcal{A}$ and all groups G/Γ, where Γ is a discrete subgroup of G contained in the center of G.*

For further information and proofs see e.g. Chapter 20 of John Lee, Introduction to smooth manifolds (1st edition).

(Version 2020.1)

Chapter 2

Symmetry

2.1 Lie group actions

2.1.A Definitions.

Definition 2.1.1 *A (smooth, left) action of a Lie group G on a manifold M is a smooth map*

$$\Psi \,:\, G \times M \;\to\; M\,, \qquad (g,m) \mapsto \Psi(g,m) \;:=\; \Psi_g(m)$$

such that the maps

$$\Psi_g : M \to M\,, \quad g \in G\,,$$

so defined have the following properties:

i. $\Psi_e = \mathrm{id}_M$

ii. $\Psi_{gh} = \Psi_g \circ \Psi_h$ *for all* $g, h \in G$.

Properties i. and ii. imply that, for each $g \in G$, Ψ_g has the smooth inverse $\Psi_{g^{-1}}$. Hence, each of the maps Ψ_g, $g \in G$, is a *diffeomorphism* of M onto itself.

Particularly in applications, there is a tendency to identify the action Ψ with the group G (even though a given group may act in different ways on a given manifold). Correspondingly, it is customary to write $g.m$ instead of $\Psi_g(m)$. We will use the two notations interchangeably.

We note that, properly, the action defined in Definition 2.1.1 is a "left action". A *right action* is defined similarly, but requiring $\Psi_{gh} = \Psi_h \circ \Psi_g$ in ii. As shown in an Exercise below, there is a one-to-one correspondence between left and right actions and—even if in applications one may need to consider both of them, for instance because a group acts on the left and on the right on a given manifold—the theory developed for one case applies also to the other.

We introduce now a few basic notions relative to an action Ψ of a Lie group G on a manifold M. First we note that, besides the maps $\Psi_g : M \to M$, it is useful to consider also the 'orbit' maps

$$\Psi^m \ : \ G \ \to \ M \, , \qquad g \mapsto \Psi^m(g) := \Psi(g, m) \, . \tag{2.1.1}$$

Next:

1. The Ψ-*orbit* of a point $m \in M$ is the set
$$\mathcal{O}_m \ := \ \{\Psi_g(m) \ : \ g \in G\} \ = \ \Psi^m(G) \ = \ G.m$$
2. The action is *transitive* if there is only one orbit: $M = \mathcal{O}_m$ for some (and hence every) point $m \in M$. Equivalently: given any two points $m, m' \in M$ there exists $g \in G$ such that $m' = \Psi_g(m)$.
3. The action is *free* if
$$g \neq e \quad \Longrightarrow \quad \Psi_g(m) \neq m \qquad \forall \, m \in M \, .$$
4. The *isotropy (sub)group* (or *stabilizer*) of a point $m \in M$ is the set of all elements of G that leave m fixed:
$$G_m := \{g \in G \ : \ \Psi_g(m) = m\} \, .$$
An action is free if and only if all its isotropy subgroups are trivial, $G_m = \{e\}$ for all $m \in M$.

Exercises 2.1.1 *(i) Show that an action is free if and only if, for any $g, h \in G$ and $m \in M$, $g.m = h.m$ implies $g = h$.*

(ii) Verify that the 'isotropy subgroups' G_m are actually subgroups of G.

(iii) Show that each isotropy subgroup G_m is a Lie subgroup of G. [Hint: $G_m = (\Psi^m)^{-1}(m)$.]

(iv) Show that the isotropy subgroups of the points of an orbit are all conjugate to each other and, if the group is abelian, they are all equal.

(v) Show that a map $\Psi : G \times M \to M$ is a (left) action of G on M if and only if the map $\tilde{\Psi} : G \times M \to M$, $(g, m) \mapsto \tilde{\Psi}(g, m) := \Psi(g^{-1}, m)$, is a right action of G on M.

2.1.B Examples. We give now a number of examples, some for illustrative purposes and others which will be used later.

Example 1. The flow Φ_t^X of a complete vector field $X \in \mathcal{X}(M)$ is an action of $\mathbb{R}$ on M. The fixed points of this action are the zeroes of X ('equilibria', in the dynamical system's terminology). If $m \in M$ is not an equilibrium point of X (i.e. $X(m) \neq 0$), it can still happen that $\Phi_T^X(m) = m$ for some $T \in \mathbb{R}$. Indeed, this happens if $\gamma(t) = \Phi_t^X(m)$ is a periodic orbit of X. So the action is free if and only if X is everywhere $\neq 0$ and X possesses no periodic orbits. All $\mathbb{R}$-actions are of this type:

Proposition 2.1.2 *Every action Ψ of $\mathbb{R}$ on a manifold M is the flow of a vector field, its "infinitesimal generator"*

$$X(m) := \frac{d}{dt}\Psi_t(m)_{|t=0}\,, \qquad m \in M\,.$$

Proof. Since $\Psi_0(m) = m$, $X(m) \in T_mM$ for all m. Furthermore, the map $m \to M$ is smooth. Hence, $X \in \mathfrak{X}(M)$. To prove that $\Psi = \Phi^X$ we must prove that, for each $m \in M$, $t \mapsto \Psi_t(m)$ is the integral curve of X through m. In fact, $\Psi_0(m) = m$ and, for all t,

$$\tfrac{d}{dt}\Psi_t(m) = \tfrac{d}{ds}\Psi_{t+s}(m)|_{s=0} = \tfrac{d}{ds}\Psi_s \circ \Psi_t(m)|_{s=0} = X(\Psi_t(m))\,. \qquad \blacksquare$$

Example 2. The 'linear action' Ψ of a Lie subgroup H of GL(n) on $\mathbb{R}^n$ is given by

$$\Psi_A(x) = Ax \qquad A \in H\,,\ x \in \mathbb{R}^n\,.$$

For instance, the linear action of SO(3) on $\mathbb{R}^3$ rotates in all possible ways each point of $\mathbb{R}^3$: $(R, x) \mapsto Rx$ for all $R \in \mathrm{SO}(3)$, $x \in \mathbb{R}^3$. The orbit of the origin $0 \in \mathbb{R}^3$ is the origin itself, and that of a point $x \neq 0$ is the sphere of radius $\|x\|$. Hence the action is neither transitive nor free. The isotropy subgroup G_x of a point $x \in \mathbb{R}^3$ is the entire SO(3) if $x = 0$ and, if $x \neq 0$, it is the subgroup of SO(3) formed by the rotations of axis x, namely $\{\exp(t\widehat{x}) : t \in \mathbb{R}\}$ (which is isomorphic to S^1).

Example 3. If an action $\Psi : G \times M \to M$ leaves a submanifold N of M invariant, namely $\Psi_g(N) \subseteq N$ for all $g \in G$, then Ψ restricts to an action

$$\Psi|_{G\times N} : G \times N \to N$$

of G on N. For instance, the linear action of SO(3) on $\mathbb{R}^3$ restricts to an action of SO(3) on the unit sphere of $\mathbb{R}^3$. This restricted action is now transitive, but still not free.

Example 4. Any Lie group G acts on itself by left-translations: the map

$$L : G \times G \to G\,, \qquad (g, h) \mapsto L_g(h) = gh\,,$$

is an action. It is free ($L_gh = gh = h$ implies $g = hh^{-1} = e$) and transitive (for any $h, h' \in G$ there exists $g \in G$ such that $h' = gh$: $g = h'h^{-1}$).

Example 5. A Lie group acts on itself also by right-translations. However, the map $R : G \times G \to G$ defined by $R_g(h) = hg$ is a right action. A left action of G on itself by right-translations is given by $(g, h) \to R_{g^{-1}}h = hg^{-1}$ (see Exercise 2.1.1.v).

Example 6. The action of a Lie group G on itself by conjugation is

$$C : G \times G \to G\,, \qquad (g,h) \mapsto C_g h \;=\; ghg^{-1}\,.$$

Its orbits are called *conjugacy classes*. This action is not transitive (it has the fixed point e) and not free (the isotropy subgroup of e is the entire group). The isotropy subgroup of an element $g \in G$ is the subgroup

$$G_h = \{g \in G :\; gh = hg\}$$

and is called the *centralizer* of h.

Example 7. Consider a vector field $X \in \mathfrak{X}(M)$ with periodic flow, namely, all its integral curves are periodic (or constant). We ask if its flow defines an action of $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$ as well. Define the 'minimal period function' $T : M \to \mathbb{R}$ of the flow of X as follows: if m is an equilibrium, then $T(m) = 0$. If not, $T(m) > 0$ is defined by $\Phi^X_{T(m)}(m) = m$ and $\Phi^X_t(m) \neq m$ for $0 < t < T(m)$. Obviously T is constant along the orbits of X (it is a 'first integral' of X), but it might not be a smooth function (even though in most cases it is smooth). If T is smooth, then the map

$$\Psi : \mathbb{T}^1 \times M \to M\,, \qquad (\alpha\,\mathrm{mod}1, m) \mapsto \Psi_{\alpha\,\mathrm{mod}1}(m) := \Phi^X_{\alpha T(m)}(m)$$

is well defined ($\Phi^X_{\alpha T(m)}(m)$ does not depend on the choice of $\alpha \in \mathbb{R}$ in the equivalence class $\alpha\,\mathrm{mod}1$ because $\Phi^X_{kT(m)}(m) = m$ for all $k \in \mathbb{Z}$) and, as is easily checked, is an action of $\mathbb{T}^1$.

Example 8. Let $\Psi^G : G \times M \to M$ and $\Psi^H : H \times M \to M$ be two actions of two Lie groups G and H on a manifold M. We ask if the map

$$\Psi^{G\times H} : (G \times H) \times M \to M\,, \qquad \Psi^{G\times H}_{(g,h)}(m) := \Psi^G_g \circ \Psi^H_h(m)$$

obtained by 'composing' the two actions is an action of $G \times H$ on M. Smoothness is obvious and, since the identity element $e_{G\times H}$ of $G \times H$ is (e_G, e_H), so is $\Psi^{G\times H}_{e_{G\times H}} = \mathrm{id}_M$. However,

$$\Psi^{G\times H}_{(g_1,h_1)(g_2,h_2)} = \Psi^{G\times H}_{(g_1g_2,h_1h_2)} = \Psi^G_{g_1g_2} \circ \Psi^H_{h_1h_2} = \Psi^G_{g_1} \circ \Psi^G_{g_2} \circ \Psi^H_{h_1} \circ \Psi^H_{h_2}$$

and

$$\Psi^{G\times H}_{(g_1,h_1)} \circ \Psi^{G\times H}_{(g_2,h_2)} = \Psi^G_{g_1} \circ \Psi^H_{h_1} \circ \Psi^G_{g_2} \circ \Psi^H_{h_2}$$

coincide for all $g_1, h_1 \in G_1$ and $g_2, h_2 \in G_2$ if and only if the two actions commute, namely

$$\Psi^G_g \circ \Psi^H_h = \Psi^H_h \circ \Psi^G_g \qquad \forall\, g \in G\,,\; h \in H$$

(for the 'only if' part, take $g_1 = e_G$, $h_2 = e_H$).

Example 9. By the previous example and Proposition A.2.7, if X_1 and X_2 are two commuting vector fields on a manifold M, $[X_1, X_2] = 0$, then

$$\Phi^{X_1,X_2} : \mathbb{R}^2 \times M \to M\,, \qquad \Phi^{X_1,X_2}_{t_1,t_2}(m) := \Phi^{X_1}_{t_1} \circ \Phi^{X_2}_{t_2}(m)$$

is an action of $\mathbb{R}^2$ on M.

Example 10. The *adjoint action* of a Lie group G on its Lie algebra $\mathfrak{g}$ is defined as follows. For any $g \in G$, the conjugation $C_g : G \to G$ is a Lie group isomorphism. Therefore, its tangent map $T_eC_g : \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra isomorphism. The map

$$\mathrm{Ad} : G \to \mathrm{Isom}(\mathfrak{g})\,, \qquad g \mapsto \mathrm{Ad}_g := T_eC_g$$

is called the *adjoint map* of G. The *adjoint action*, or *adjoint representation*, of G is the map

$$\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}\,, \qquad (g, \xi) \mapsto \mathrm{Ad}_g\xi\,.$$

That it is an action of G on $\mathfrak{g}$ is verified observing that, since $C_e = \mathrm{id}_G$, $\mathrm{Ad}_e\xi = T_e(\mathrm{id}_G) \cdot \xi = \xi$ for all $\xi \in \mathfrak{g}$ and, using the chain rule and $C_{g'}(e) = e$, $\mathrm{Ad}_{gg'} = T_e(C_{gg'}) = T_e(C_g \circ C_{g'}) = T_{C_{g'}(e)}C_g \circ T_eC_{g'} = T_eC_g \circ T_eC_{g'} = \mathrm{Ad}_g \circ \mathrm{Ad}_{g'}$ for all $g, g' \in G$.

Note that the fact that the adjoint map is a Lie algebra isomorphism implies

$$\begin{aligned} C_g \circ \exp_G \;&=\; \exp_G \circ \mathrm{Ad}_g & \forall g \in G & \qquad (2.1.2)\\ \mathrm{Ad}_g[\![\xi, \eta]\!] \;&=\; [\![\mathrm{Ad}_g\xi, \mathrm{Ad}_g\eta]\!] & \forall\, g \in G\,,\ \xi, \eta \in \mathfrak{g}\,. & \qquad (2.1.3)\end{aligned}$$

Example 11. Adjoint representation of GL(n). For $A, B \in \mathrm{GL(n)}$, $C_A(B) = ABA^{-1}$. For $A \in \mathrm{GL(n)}$ and $V \in T_A\mathrm{GL(n)} = \mathrm{L}(n)$, $t \mapsto \exp(At)$ is a curve in GL(n) through A. Thus

$$Ad_AV = T_{\mathbb{I}}C_A \cdot V = \tfrac{d}{dt}C_A(\exp(tV))|_{t=0} = \tfrac{d}{dt}A\exp(tV)A^{-1}|_{t=0} = AVA^{-1}\,.$$

We conclude that the adjoint representation of GL(n) is the action of GL(n) on $\mathrm{L}(n)$ given by $\mathrm{Ad}_AV := AVA^{-1}$. Note that the identities (2.1.2) and (2.1.3) become the familiar identities $A\exp(V)A^{-1} = \exp(AVA^{-1})$ and $A[U, V]_-A^{-1} = [AUA^{-1}, AVA^{-1}]_-$.

Example 12. The *coadjoint action*, or *coadjoint representation*, of a Lie group G is the action Ad^* of G on the dual $\mathfrak{g}^*$ of $\mathfrak{g}$ obtained by dualizing the adjoint action.[1] Specifically, it is defined as

$$\mathrm{Ad}^* \,:\, G \times \mathfrak{g}^* \to \mathfrak{g}^*\,, \qquad (g,\mu) \mapsto (\mathrm{Ad}^*)_g \mu := (Ad_{g^{-1}})^* \mu\,. \tag{2.1.4}$$

(Note the different meanings of $*$ in these formulas: in Ad^*, it denotes the name of the coadjoint action, and $(\mathrm{Ad}^*)_g : \mathfrak{g} \to \mathfrak{g}$ is the map at g fixed of this action; instead, $(\mathrm{Ad}_g)^*$ is the adjoint of the map Ad_g). The reason why $(\mathrm{Ad}^*)_g$ is defined as the adjoint of $\mathrm{Ad}_{g^{-1}}$ and not of Ad_g is because this gives a left action instead of a right action. That Ad^* is an action is easily verified: $(\mathrm{Ad}^*)_{gh} = (\mathrm{Ad}_{(gh)^{-1}})^* = (\mathrm{Ad}_{h^{-1}} \circ \mathrm{Ad}_{g^{-1}})^* = (\mathrm{Ad}_{g^{-1}})^* \circ (\mathrm{Ad}_{h^{-1}})^*$.

Example 13. The expressions of the adjoint and coadjoint representations of SO(3) depend, of course, on the identification of the Lie algebra $\mathfrak{so}(3)$ with either $(\mathrm{skew}(3), [\ ,\ ]_-)$ or $(\mathbb{R}^3, \times)$. In the first case we may write the matrices of skew(3) as $\hat{v}$ with $v \in \mathbb{R}^3$. Then, from example 11. we know that $\mathrm{Ad} : \mathrm{SO}(3) \times \mathrm{skew}(3) \to \mathrm{skew}(3)$ is given by

$$\mathrm{Ad}_R \hat{v} = R\hat{v}R^{-1}\,.$$

Given that $R\hat{v}R^{-1} = \widehat{Rv}$ and that the Lie algebra isomorphism between the two is the hat map $\hat{}: \mathbb{R}^3 \to \mathrm{skew}(3)$, it follows that, if the identification of $\mathfrak{so}(3)$ with $\mathbb{R}^3$ is used, then

$$\mathrm{Ad}_R v = Rv\,.$$

In order to express the coadjoint representation of SO(3) we need to choose an inner product on $\mathfrak{so}(3)$, so as to identify $\mathfrak{so}(3)^*$ and $\mathfrak{so}(3)$. If we identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$, then we may use the Euclidean inner product on $\mathbb{R}^3$. The matrix of the adjoint of a linear map is then the transpose of the matrix of the linear map and so $(Ad_R)^* = R^*$. Thus

$$(Ad_R)^* v = R^T v$$

and

$$(\mathrm{Ad}^*)_R v = (\mathrm{Ad}_{R^{-1}})^* v = Rv\,.$$

14. Lifted actions. Each action Ψ of a Lie group G on a manifold M induces an action of G on the tangent bundle of M, its *tangent lift* Ψ^{TM}. Lifted actions play an important role in Lagrangian mechanics, because Lagrange equations are second order equations: as such, they can be viewed as a vector field on the tangent bundle TM of the 'configuration space' M, and it is natural to 'lift' to TM (namely, to positions and velocities) symmetry transformations defined in M (namely, acting on positions).

[1] Recall that the adjoint of a linear map $\ell : E \to F$ between two vector spaces E and F is the map $\ell^* : F^* \to E^*$ between their duals defined as follows: for any $\alpha \in F^*$, $\ell^*(\alpha) \in E^*$ is the linear map $E \to \mathbb{R}$ such that $\langle \ell^*(\alpha), v\rangle_E = \langle \alpha, \ell(v)\rangle_F$ for all $v \in E$.

Definition 2.1.3 *The* lifted action *(or tangent lift) of an action Ψ of a Lie group G on a manifold M is the map*

$$\Psi^{TM} : G \times TM \to TM\,, \qquad (g,v) \mapsto \Psi_g^{TM}(v)$$

with

$$\Psi_g^{TM} := T\Psi_g \qquad \forall\, g \in G\,.$$

Explicitly,[2] $\Psi_g^{TM}(v) = T_m\Psi_g \cdot v$ or $\Psi_g^{TM}(m,v) = \big(\Psi_g(m), T_m\Psi_g \cdot v\big)$ for all $g \in G$, $m \in M$, $v \in T_mM$. In bundle coordinates $(x,\dot x)$ of TM,

$$(\Psi_g^{TM})^{\rm loc}(x,\dot x) = \Big(\Psi_g^{\rm loc}(x), (\Psi_g^{\rm loc})'(x)\dot x\Big)\,.$$

Proposition 2.1.4 *If Ψ is an action of G on M, then Ψ^{TM} is an action of G on TM.*

Proof. $\Psi_e^{TM} = {\rm id}_{TM}$ because $\Psi_e = {\rm id}_M$. For any $g,h \in G$, $\Psi_{gh}^{TM} = T\Psi_{gh} = T(\Psi_g \circ \Psi_h) = (T\Psi_g) \circ (T\Psi_h) = \Psi_g^{TM} \circ \Psi_h^{TM}$. ■

Examples: 1. For any $u \in \mathbb{R}^n$, the lift of the action $\lambda.x = x + \lambda u$ of $\mathbb{R} \ni \lambda$ on $\mathbb{R}^n \ni x$ (translations parallel to u) is the action of $\mathbb{R}$ on $T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n \ni (x,v)$ given by $\lambda.(x,v) = (x+\lambda u, v)$.

2. The lift of the linear action of $GL(n) \ni A$ on $\mathbb{R}^n \ni x$, given by $A.x = Ax$, is $A.(x,v) = (Ax, Av)$.

Exercises 2.1.2 *(i) The computation of $\frac{d}{dt}\Psi_t(m)$ in the proof of Proposition 2.1.2 is based on writing Ψ_{t+s} as $\Psi_s \circ \Psi_t$. However, Ψ_{t+s} could also be written as $\Psi_s \circ \Psi_t$ (as is indeed done for $\gamma(t+s)$ in the—under other respects equal—proof of Proposition 1.3.18) and this would lead to a different expression for $\frac{d}{dt}\Psi_t(m)$. Determine this expression and then show that,* after it has been proved that *$\Psi = \Phi^X$, it can be shown that the two expressions are equal. [Hint: Since $[X,X]=0$, $(\Phi_t^X)_*X = X$ for all t.]*

(ii) Why the lift of an action Ψ in not defined as $T\Psi$?

[2] The points of TM are the vectors tangent to M at its points. To keep track of the base point, we may denote them

- either as vectors $v \in T_mM$, or perhaps better $v_m \in T_mM$, with $m \in M$
- or as pairs (m,v) with $m \in M$ and $v \in T_mM$.

When we write the value of a vector field $X : M \to TM$ at a point $m \in M$ as $X(m) \in T_mM$ we are adopting the first possibility. When useful, we turn to the second.

Remark: Since the set $\mathrm{Diff}(M)$ of all diffeomorphisms on a manifold forms a group under composition, an action Ψ of G on M may also be regarded as a group homomorphism $\widehat{\Psi}$ between G and $\mathrm{Diff}(M)$, which to any element g of G associates the diffeomorphism Ψ_g of G. An action by linear maps of a Lie group G on a vector space E is called a *representation* of G on E and can be regarded as a Lie group homomorphism $\widehat{\Psi}$ between G and $\mathrm{Isom}(E)$, the group of all linear isomorphisms of E. A representation is *faithfully* if the map $\widehat{\Psi}$ is injective. A faithful representation $\widehat{\Psi}$ is thus a Lie group isomorphism between G and $\widehat{\Psi}(G) \subseteq \mathrm{Isom}(E)$, and gives a realization of G by means of a group of linear transformations of a vector space (hence, as a group of matrices; not all Lie groups admit faithfully representations in $\mathrm{GL}(n, \mathbb{R})$, though; a counterexample is $\mathrm{SL}(2, \mathbb{R})$).

2.1.C Invariance and equivariance.

Definition 2.1.5 *Consider a Lie group G and a map $\mathcal{F} : N \to P$ between two manifolds N and P.*

$$\begin{tikzcd} N \arrow[d, "\Psi_g"'] \arrow[dr, "\mathcal{F}"] & \\ N \arrow[r, "\mathcal{F}"] & P \end{tikzcd}$$

i. $\mathcal{F}$ is invariant *under an action Ψ of G on N if*

$$\mathcal{F} \circ \Psi_g = \mathcal{F} \qquad \forall g \in G\,. \tag{2.1.5}$$

$$\begin{tikzcd} N \arrow[r, "\mathcal{F}"] \arrow[d, "\Psi^N_g"'] & P \arrow[d, "\Psi^P_g"] \\ N \arrow[r, "\mathcal{F}"] & P \end{tikzcd}$$

ii. $\mathcal{F}$ is equivariant *with respect to an action Ψ^N of G on N and to an action Ψ^P of G on P if*

$$\mathcal{F} \circ \Psi^N_g = \Psi^P_g \circ \mathcal{F} \qquad \forall g \in G\,. \tag{2.1.6}$$

Proposition 2.1.6 *Assume that $\mathcal{F} : N \to P$ is equivariant with respect to two actions Ψ^N on N and Ψ^P on P of a Lie group G. Then, if Ψ^N is transitive, $\mathcal{F}$ has constant rank.*

Proof. Differentiating the left- and right-hand sides of (2.1.6) at a point $n \in N$ gives

$$T_{\Psi^N_g(n)}\mathcal{F} \cdot T_n\Psi^N_g \;=\; T_{\mathcal{F}(n)}\Psi^P_g \cdot T_n\mathcal{F} \qquad \forall g \in G\,. \tag{2.1.7}$$

Since Ψ^N_g and Ψ^P_g are diffeomorphisms, the linear maps $T_n\Psi^N_g$ and $T_{\mathcal{F}(n)}\Psi^P_g$ are isomorphisms and hence, for any $n \in N$ and $g \in G$, $T_{\Psi^N_g(n)}\mathcal{F}$ has the same rank as $T_n\mathcal{F}$. This proves that the rank of $\mathcal{F}$ is constant on each orbit of Ψ^N. The statement now follows from the fact that, if Ψ^N is transitive, then N is an orbit. ■

$$\begin{tikzcd} G \arrow[r, "\mathcal{F}"] \arrow[d, "L_g"'] & H \arrow[d, "L_{\mathcal{F}(g)}"] \\ G \arrow[r, "\mathcal{F}"] & H \end{tikzcd}$$

Exercises 2.1.3 *(i) Interpret invariance as a particular case of equivariance.*

(ii) Show that a Lie group homomorphism $\mathcal{F} : G \to H$ is equivariant with respect to a transitive action of G on G and an action of G on H, so as to deduce statement ii of Proposition 1.1.7 from Proposition 2.1.6. [Hint: $\Psi^H : (g, h) = L_{\mathcal{F}(g)}h$ is an action of G on H].

2.1.D Actions of one-parameter subgroups. One-parameter subgroups of a Lie group G have been introduced in section 1.3.C. Restricting an action of G to these subgroups give actions of $\mathbb{R}$ generated by the vectors in the algebra, which turn out to play an useful role in investigating some properties of the action.

Definition 2.1.7 *Let Ψ be an action of a Lie group G on a manifold M. The* infinitesimal generator of the action associated to a vector $\xi \in \mathfrak{g}$ *is the vector field ξ^M on M defined by*

$$\xi^M(m) := \tfrac{d}{dt}\Psi_{\exp(t\xi)}(m)|_{t=0} \qquad \forall m \in M\,. \tag{2.1.8}$$

Note that (2.1.8) actually defines a vector field on M because $\exp(t\xi)(m) = m$ if $t = 0$.

Note also that (2.1.8) gives

$$\xi^M(m) = T_e\Psi^m \cdot \xi \qquad \forall \xi \in \mathfrak{g}, m \in M\,. \tag{2.1.9}$$

Furthermore, ξ^M is the infinitesimal generator, in the sense of Proposition 2.1.2, of the $\mathbb{R}$-action of the one parameter subgroup $\{\Psi_{\exp(t\xi)} : t \in \mathbb{R}\}$ of G. Therefore, the flow of ξ^M coincides with this action:

$$\Phi_t^{\xi^M} = \Psi_{\exp(t\xi)} \qquad \forall\, t \in \mathbb{R}\,.$$

This has the following consequence:

Proposition 2.1.8 *If an action is free, then all its infinitesimal generators ξ^M, $\xi \neq 0$, are everywhere nonzero in M.*

Proof. If $\xi^M(\overline{m}) = 0$ for a certain $\xi \in \mathfrak{g} \setminus \{0\}$ and a certain $\overline{m} \in M$, then $\overline{m}$ is an equilibrium of ξ^M and therefore $\Psi_{\exp(t\xi)}(\overline{m}) = \overline{m}$ for all t. This contradicts the freeness of Ψ because, since exp is a local diffeomorphism at 0, there is some t for which $\exp(t\xi) \neq e$. ■

We note two further properties of these vector fields:

Proposition 2.1.9 *For all $\xi, \eta \in \mathfrak{g}$:*

i. $(\Psi_g)_*\xi^M = (Ad_g\xi)^M$

ii. $[\xi^M, \eta^M] = -[\![\xi, \eta]\!]^M$

Proof. (i) First note that, for any $m \in M$ and $g \in G$,

$$\begin{aligned}((\Psi_g)_*\xi^M)(\Psi_g(m)) \;&=\; (T\Psi_g \cdot \xi^M)(m) \;=\; T_m\Psi_g \cdot \tfrac{d}{dt}\Psi_{\exp(t\xi)}(m)|_{t=0} \\ &=\; \tfrac{d}{dt}\Psi_g \circ \Psi_{\exp(t\xi)}(m)|_{t=0} \;=\; \tfrac{d}{dt}\Psi_{g\exp(t\xi)}(m)|_{t=0}\,.\end{aligned}$$

Evaluating this equality at $\Psi_g^{-1}(m)$ instead of m, and recalling the definitions 2.1.1 of the orbit map Ψ^m and of the Adjoint map (example 10 in section 2.1.B), gives

$$\begin{aligned} ((\Psi_g)_*\xi^M)(m) &= \tfrac{d}{dt}\Psi_{g\exp(t\xi)}(\Psi_{g^{-1}}(m))|_{t=0} = \tfrac{d}{dt}\Psi_{g\exp(t\xi)g^{-1}}(m)|_{t=0} \\ &= \tfrac{d}{dt}\Psi_{C_g(\exp(t\xi))}(m)|_{t=0} = \tfrac{d}{dt}\Psi^m(C_g(\exp(t\xi))|_{t=0} \\ &= T_{C_g(e)}\Psi^m \cdot T_eC_g \cdot \xi = T_e\Psi^m \cdot \mathrm{Ad}_g(\xi) \\ &= (\mathrm{Ad}_g(\xi))^M . \end{aligned}$$

(ii) This is omitted. ■

Exercises 2.1.4 *(i) We have associated two different vector fields to a Lie algebra vector ξ: the left-invariant vector field X_ξ on G and the infinitesimal generator of the action ξ^M on M. Show that*

$$T_g\Psi^m \cdot X_\xi(g) = T_m\Psi_g \cdot \xi^M(m) \qquad \forall m \in M,\, g \in G\,.$$

[Hints: First verify that, for any m, $\Psi^m : G \to M$ is equivariant with respect to the action of G by left-translation on itself and to the action Ψ on M, namely $\Psi_g \circ \Psi^m(h) = \Psi^m \circ L_g(h)$ for all $g, h \in G$. Then uses (2.1.7).]

2.1.E Structure of the orbits. Even though in the next section we will give, under stronger hypotheses, stronger results, it may be useful to say something about the structure of the orbits in the general case.

Proposition 2.1.10 *The orbits of a free action Ψ of a Lie group G on a manifold M are immersed submanifolds of M diffeomorphic to G.*

Proof. The orbit $\mathcal{O}_m = \{\Psi_g(m) : g \in G\}$ of a point m is the image of the map $\Psi^m : G \to M$, and we show that this map is an injective immersion.

Since $\Psi^m(g) = \Psi^m(h)$ is equivalent to $\Psi_g(m) = \Psi_h(m)$, injectivity follows from the freeness of the action.

Ψ^m is an immersion at e if $T_e\Psi^m : T_eG \to T_mM$ is injective, namely if $T_e\Psi^m \cdot \xi \neq 0$ for all nonzero $\xi \in \mathfrak{g}$. By (2.1.9), $T_e\Psi^m \cdot \xi = \xi^M(m)$ which, according to Proposition 2.1.8, is nonzero.

The map Ψ^m is equivariant with respect to the (transitive) action of G by left-translation on itself and to the action Ψ on M. Indeed, $\Psi_g \circ \Psi^m(h) = \Psi_g \circ \Psi_h(m) = \Psi_{gh}(m) = \Psi_{L_gh}(m) = \Psi^m(L_g(h))$ for all $g, h \in G$.

Thus, by Proposition 2.1.6, the fact that Ψ is immersive at e implies that it is immersive at all points of G.

Finally, an injective immersion is a diffeomorphism onto its image. ■

Exercises 2.1.5 *(i) Show that, if Ψ is free, then $T_m\mathcal{O}_m = \{\xi^M(m) : \xi \in \mathfrak{g}\}$. [Hint: if $\iota : P \to Q$ is an immersion, then $T_\iota f(T_pP) = T_{\iota(p)}\iota(P)$ for all $p \in P$.]*

2.2 Reduction of invariant vector fields

2.2.A Invariant vector fields.

Definition 2.2.1 *A vector field $X \in \mathfrak{X}(M)$ is* invariant *under an action Ψ of a Lie group G on a manifold M if*

$$(\Psi_g)_* X = X \qquad \forall\, g \in G\,. \tag{2.2.1}$$

In such a case, Ψ is said to be a symmetry action *of X and G a* symmetry group *of X.*

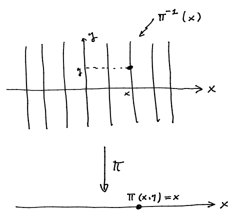
Condition (2.2.1) may be written as

$$X(\Psi_g(m)) = T_m\Psi_g \cdot X(m) \qquad \forall m \in M,\ g \in G$$

(in coordinates, $X^{\rm loc}(\Psi_g^{\rm loc}(x)) = (\Psi_g^{\rm loc})'(x)X^{\rm loc}(x)$) and shows that, for each g and m, the value of a Ψ-invariant vector field X at the point $\Psi_g(m)$ is obtained by transforming with the tangent map $T_m\Psi_g$ its values at the point m. Also, by Proposition A.1.3, condition (2.2.1) is equivalent to

$$\Psi_g \circ \Phi_t^X = \Phi_t^X \circ \Psi_g \qquad \forall t \in \mathbb{R},\ g \in G \tag{2.2.2}$$

Examples: 1. Consider the action of $\mathbb{R} \ni t$ on $\mathbb{R}^n \ni x$ by translations along a fixed direction, given by a nonzero $u \in \mathbb{R}^n$, namely $\Psi_t(x) = x + tu$. We may work in coordinates. Clearly $\Psi_t'(x) = \mathbb{I}$ for all x and the condition of invariance of a vector field X is



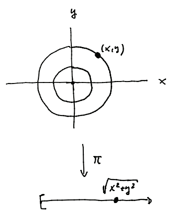
$$X(x+tv) \;=\; X(x) \qquad \forall t \in \mathbb{R}\,, x \in \mathbb{R}^n\,.$$

Thus, a vector field in $\mathbb{R}^n$ is invariant under translations along v if it is constant along each line parallel to v.

2. Consider now the linear action Ψ of SO(2) on $\mathbb{R}^2$, namely

$$\Psi_R(x) \;=\; Rx\,, \qquad R \in \mathrm{SO}(2)\,, \quad x \in \mathbb{R}^2\,.$$

A vector field X on $\mathbb{R}^2$ is invariant under this action if, for all R and x, $X(Rx) = \Psi_R'(x)X(x)$. Since Ψ_R is a linear map, $\Psi'(x) = R$ for all x and this condition is



$$X(Rx) \;=\; RX(x) \qquad \forall R, x\,.$$

Thus, for any $x \neq 0$, its value at each point Rx of the circle of radius $\|x\|$, $R \in \mathrm{SO}(2)$, is obtained by rotating with R its value at the point x. Said differently, on each circle centered at the origin, X has constant norm and forms a constant angle with the radius.

$$\begin{tikzcd} M \arrow[r, "X"] \arrow[d, "\Psi_g^M"'] & TM \arrow[d, "\Psi_g^{TM}"] \\ M \arrow[r, "X"] & TM \end{tikzcd}$$

Remark: The condition of invariance of a vector field $X \in \mathcal{X}(M)$ under an action Ψ^M on a manifold M, namely $X \circ \Psi_g^M = T_m\Psi_g \circ X$ for all $g \in G$, can be written $X \circ \Psi_g = \Psi_g^{TM} \cdot X$ for all $g \in G$ (where Ψ^{TM} is the lifted action of Ψ^M, see Definition 2.1.3). Hence, the invariance of X under an action Ψ^M is equivalent to the equivariance of X, regarded as a map $X : M \to TM$, with respect to the actions Ψ^M on M and Ψ^{TM} on TM. This explains why, in the literature, invariant vector fields are also called equivariant vector fields.

Exercises 2.2.1 *(i) Assume that $X \in \mathcal{X}(M)$ is invariant under an action Ψ on M and that $X(m^*) = 0$ for a certain $m^* \in M$. Show that the Ψ-orbit of m^* consists entirely of equilibria (= zeroes) of X.*

(ii) Let $Fix(\Psi) := \{m \in M : \Psi_g(m) = m \ \forall, g \in G\}$ be the set of fixed points of an action Ψ of a Lie group G on M. Show that $Fix(\Psi)$ is invariant under the flow of any Ψ-invariant vector field (namely, if $X \in \mathcal{X}(M)$ is Ψ-invariant, then $\Phi_t^X(Fix(\Psi)) \subseteq Fix(\Psi)$ for all $t \in \mathbb{R}$). [Hint: use (2.2.2)].

(iii) Show that, in example 2., if X is SO(2)*-invariant then $X(0) = 0$. [Hints: use either a continuity argument or the result of the previous exercise].*

2.2.B The quotient space. If G acts on M, then there is one and only one G-orbit through each point of M, and M is decomposed in the union of the G-orbits. Belonging to a G-orbit is an equivalence relation $\sim$ on M:

$$m \sim m' \quad \Leftrightarrow \quad m' \in \mathcal{O}_m \,.$$

Denote by $[m]$ the equivalence class of $m \in M$ (namely, the G-orbit through m). The set M/G of all equivalence classes is called the *quotient space* or the *orbit space*. We denote by π the *quotient map* or *canonical projection*

$$\pi : M \ \to \ M/G \,, \quad m \mapsto [m] \,.$$

The quotient space M/G is a topological space with the *quotient topology*, which is defined as follows:

$$U \subseteq M/G \quad \text{is open} \quad \Leftrightarrow \quad \pi^{-1}(U) \subseteq M \quad \text{is open} \,.$$

If M/G is equipped with the quotient topology, then the canonical projection $\pi : M \to M/G$ is continuous.[3] However, the question of whether M/G is *a smooth manifold* is subtler. Before investigating this question, we note that:

Proposition 2.2.2 *Assume that $\overline{M} := M/G$ has a smooth structure such that the quotient map $\pi : M \to M/G$ is a submersion. Then, for any Ψ-invariant vector field $X \in \mathcal{X}(M)$ there exists a vector field $\overline{X} \in \mathcal{X}(\overline{M})$ which is π-related to X and hence*

$$\Phi_t^{\overline{X}} \circ \pi = \pi \circ \Phi_t^X \qquad \forall t \in \mathbb{R} \,. \tag{2.2.3}$$

[3]The quotient topology is the final topology induced by π on M/G: the finest topology that makes π continuous.

Proof. First, fix $t \in \mathbb{R}$. By (2.2.2), Φ_t^X maps the entire G-orbit of a point $m \in M$ into another G-orbit. Thus, $\pi(\Phi_t^X(m)) \in \overline{M}$ assumes the same value in all points m belonging to a G-orbit, and for any $\overline{m} \in \overline{M}$ we may *define* $\overline{\Phi_t^X}(\overline{m})$ as $\pi(\Phi_t^X(m))$ with any $m \in \pi^{-1}(\overline{m})$. Since π is surjective, this defines a map

$$\overline{\Phi_t^X} : \overline{M} \to \overline{M}$$

which by its very construction satisfies $\overline{\Phi_t^X} \circ \pi = \pi \circ \Phi_t^X$ for all t.

Putting together all these maps gives a map

$$\overline{\Phi^X} : \mathbb{R} \times \overline{M} \to \overline{M}\,, \qquad (t, \overline{m}) \mapsto (\overline{\Phi^X})_t(\overline{m})$$

with

$$(\overline{\Phi^X})_t := \overline{\Phi_t^X} \qquad \forall t\,.$$

It is not difficult to prove that $\overline{\Phi^X}$ is smooth (see the exercises). Moreover, it inherits from Φ^X the property of being an action of $\mathbb{R}$ on $\overline{M}$: for all $m \in M$ and $t, s \in \mathbb{R}$

$$(\overline{\Phi^X})_0(\pi(m)) = \overline{\Phi_0^X}(\pi(m)) = \pi \circ \Phi_0^X(m) = \pi(m)$$

and

$$\begin{aligned}
(\overline{\Phi^X})_{t+s}(\pi(m)) &= \overline{\Phi_{t+s}^X}(\pi(m)) \;=\; \pi \circ \Phi_{t+s}^X(m) \;=\; \pi \circ \Phi_t^X \circ \Phi_s^X(m) \\
&= \overline{\Phi_t^X} \circ \pi \circ \Phi_s^X(m) \;=\; \overline{\Phi_t^X} \circ \overline{\Phi_s^X} \circ \pi(m) \\
&= (\overline{\Phi^X})_t \circ (\overline{\Phi^X})_s(\pi(m))
\end{aligned}$$

(since π is surjective, checking a property at all points $\pi(m)$ with $m \in M$ is the same as checking it at all points of $\overline{M}$).

The proof is concluded by recalling that a smooth $\mathbb{R}$-action is the flow of a vector field (Proposition 2.1.2). Thus, there exists a vector field $\overline{X}$ on $\overline{M}$ such that $\overline{\Phi^X} = \Phi^{\overline{X}}$. Its flow satisfies (2.2.3) and, by Lemma (1.2.11),[4] it is π-related to X. ■

Under the hypotheses of Proposition 2.2.2 on the action Ψ, a Ψ-invariant vector field X on M is π-related to a vector field $\overline{X}$ on $\overline{M}$. Recall that this means that

$$\overline{X}(\pi(m)) = T_m\pi \cdot X(m) \qquad \forall, m \in M$$

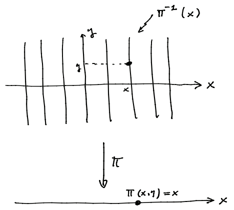
and that the flows of X and $\overline{X}$ are related by (2.2.3).

The passage from $X \in \mathfrak{X}(M)$ to $\overline{X} \in \mathfrak{X}(\overline{M})$ is called *reduction*, and $\overline{X}$ is called the *reduced* vector field of X. If, adopting the Dynamical System

[4]In fact, statement ii. of that Lemma gives the opposite implication to the one we need here. But, as its proof shows, that statement could be replaced by the following one"$Y \in \mathfrak{X}(N)$ is Ψ-related to $X \in \mathfrak{X}(M)$ if and only if $\Phi_t^Y \circ \Psi = \Psi \circ \Phi_t^X$ $\forall t$."

terminology, M is called the phase space of X, then $\overline{M}$ is called the reduced phase space.

Since the dimension of $\overline{M}$ is lower of that of M, $\dim \overline{M} = \dim M - \dim G$ if π is a submersion, there is the hope of being able to understand something about the dynamics (= the properties of the flow) of $\overline{X}$. The next question is if it is then possible to "reconstruct" the dynamics of X from that of $\overline{X}$. In order to answer this question we need to understand better the structure of the G-orbits inside M.



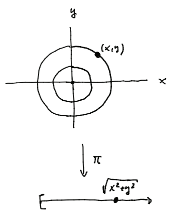
Examples: 1. Consider the action of $\mathbb{R} \ni \lambda$ on $\mathbb{R}^2 \ni (x, y)$ by translations along the y-axis:

$$\Psi_\lambda(x, y) \;=\; (x, y + \lambda)\,.$$

The orbits are the lines parallel to the y-axis. Each of them is determined by its x-coordinate. The quotient space can thus be identified with the x-axis, in the precise sense that it is in bijective correspondence with $\mathbb{R}$ with quotient map

$$\pi : (x, y) \mapsto x\,.$$

The quotient topology on $\mathbb{R}^2/\mathbb{R} = \mathbb{R}$ is the standard topology of $\mathbb{R}$ ($\pi^{-1}(U) = U \times \mathbb{R} \subset \mathbb{R} \times \mathbb{R}$ is open if and only if $U \subset \mathbb{R}$ is open). In this case, the quotient space is a smooth manifold and π is a submersion.



2. If $G = S^1$ acts by rotation of $\mathbb{R}^2$ about the origin, then its orbits are the circles centered at the origin and the origin itself. The peculiarity of the orbit of the origin is related to the fact that the isotropy subgroup of the origin is nontrivial, $G_0 = S^1$. The quotient space can be identified with the half-line $[0, \infty)$, with canonical projection $\pi : \mathbb{R}^2 \to \mathbb{R}^2/S^1$ given by

$$(x, y) \mapsto \sqrt{x^2 + y^2}\,.$$

A basis for the quotient topology is formed by the open intervals (a, b) with $b > a > 0$, whose preimages are open annuli in $\mathbb{R}^2$, and the half-open intervals $[0, b)$ with $b > 0$, whose preimages are open disks in $\mathbb{R}^2$. The quotient space has no smooth structure compatible with this topology (it is a manifold with a boundary).

3. Consider the action of $\mathbb{R} \ni \lambda$ on $\mathbb{T}^2 \ni (\alpha_1, \alpha_2)$ given by $\Psi_\lambda(\alpha_1, \alpha_2) = (\alpha_1 + \lambda\omega_1, \alpha_2 + \lambda\omega_2) \pmod 1$, with $\omega_1/\omega_2 \notin \mathbb{Q}$ (the irrational flow on $\mathbb{T}^2$). Every orbit is dense in $\mathbb{T}^2$ and the quotient topology on $\mathbb{T}^2/\mathbb{R}$ is the trivial topology, in which the open sets are the empty set and the entire space. (If $U \neq \emptyset$ is an open subset of $\mathbb{T}^2/\mathbb{R}$, then $\pi^{-1}(U)$ is an open subset of $\mathbb{T}^2$. As such, it has nonempty intersection with *any* orbit. Therefore, being a union of orbits, it is the union of *all* orbits, namely it is the entire $\mathbb{T}^2$. Thus $U = \mathbb{T}^2/\mathbb{R}$). No smooth structure is compatible with this topology.

In the second example, the obstruction to the smoothness of the quotient space comes from the non-triviality of the isotropy, and is in a way localized where the isotropy changes. Requiring the freeness of the action avoids this problem. In the third example the action is free, but the problem arises from the fact that the orbits 'keep coming back' with the result of self-accumulating on themselves. The notion of 'properness' of an action is meant precisely to prevent this behaviour:

Definition 2.2.3 *An action Ψ of a Lie group G on a manifold M is* proper *if for any compact subset $P \subset M$, the subset $\{g \in G \,:\, P \cap \Psi_g(P) \neq \emptyset\}$ of G is compact.*

It is important to keep in mind that:

Proposition 2.2.4 *Any action of a compact Lie group is proper.*

Proof. See the exercises below. ■

The main result, which we do not prove here because the proof is elaborated and rather technical, is the following:

Proposition 2.2.5 *Let Ψ be a free and proper action of a k-dimensional Lie group on an n-dimensional manifold M ($n \geq k$). Then M/G, equipped with the quotient topology, has a (unique) smooth structure such that $\pi : M \to M/G$ is smooth and is a submersion.*

In the sequel we will consider only free and proper actions. We will write $\overline{M}$ to mean M/G with this smooth structure. Note that $\dim \overline{M} = n - k$ and that the orbits are embedded submanifolds.

Exercises 2.2.2 *(i) Determine the quotient space M/G in the following cases: (a) Action $\Psi(t,(\alpha_1,\alpha_2)) = (\alpha_1 + t, \alpha_2) \bmod 2\pi$ of S^1 on $\mathbb{T}^2$. (b) Action $\Psi(t,x) \mapsto x + tv$, with $v \neq 0$ a given vector of $\mathbb{R}^2$, of $\mathbb{R}$ on $\mathbb{R}^2$. (c) Action of S^1 on $S^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$ by rotations about the x_3-axis. (d) Linear action of $\mathrm{SO}(3)$ on $\mathbb{R}^3$.*

(ii) Show that the action of $\mathbb{R}$ on $\mathbb{R}^n$ by translations along a given direction is a proper action.

(iii) Show that the orbits of a free and proper action are embedded submanifolds. [Hints: the fibers of a submersion...]

(iv) Show that the map $\overline{\Phi^X}$ defined in the proof of Proposition 2.2.2 is smooth. [Hints: Being a local question, this can be proven using coordinates. Prove that if $N \subseteq \mathbb{R}^n$ and $P \subseteq \mathbb{R}^k$ are open sets, $\psi : N \to P$ is a surjective submersion, $f : N \to N$ is a smooth map and $g : P \to P$ is a map such that $\Psi \circ f = g \circ \Psi$, then g is smooth. To do so, use the submersion theorem, according to which in a neighbourhood of any point of N there is a coordinate system in which the first k coordinates coincide with the components of ψ.]

$$\begin{tikzcd} N \arrow[r, "f"] \arrow[d, "\psi"'] & N \arrow[d, "\psi"] \\ P \arrow[r, "g"] & P \end{tikzcd}$$

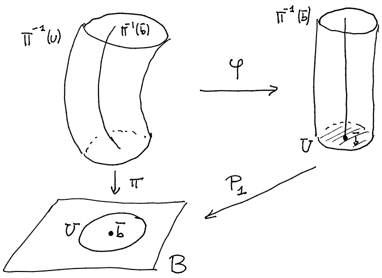
(v) Show that the maps $\overline{\Phi^X_t} : M/G \to M/G$ and $\overline{\Phi^X} : \mathbb{R} \times M/G \to M/G$ defined in the proof of Proposition 2.2.2 (but not $\overline{X}$ and $\Phi^{\overline{X}}$) can be defined even if M/G is only a topological space and $\pi : M \to M/G$ a continuous map, and that they are continuous.

(vi) A map between two topological spaces is said to be proper *if the preimages of compact sets are compact. Show that an action $\Psi : G \times M \to M$ is proper if and only if the map $\tilde{\Psi} : G \times M \to M \times M$, $(g,m) \mapsto (\Psi_g(m), m)$ is proper.*

(vii) Prove Proposition 2.2.4. [Hint: Use the previous exercise, together with the facts that the preimage of a compact set under a continuous map is closed and closed subsets of a compact space are compact.]

(viii) Prove that no free action of a non-compact Lie group on a compact manifold M is proper. [Hint: take $P = M$ in the definition.]

2.2.C Fibrations.



Definition 2.2.6 *A* locally trivial fibration *(or* fiber bundle*)* (M, B, π) *is formed by two (smooth) manifolds M and B and a surjective submersion $\pi : M \to B$ such that, for each $\overline{b} \in B$ there exist*
i. a neighbourhood U of $\overline{b}$ in B
i. a diffeomorphism $\varphi : \pi^{-1}(U) \to U \times \pi^{-1}(\overline{b})$
with the property that

$$\pi|_{U\times\pi^{-1}(U)} = p_1 \circ \varphi \tag{2.2.4}$$

where $p_1 : U \times \pi^{-1}(\overline{b}) \to \pi^{-1}(\overline{b})$, $(b, y) \mapsto b$, is the projection onto the first factor.

M is called the total space*, B the* base *and the maps φ the* local trivializations *of (M, B, π).*

Condition (2.2.4) is equivalent to saying that each local trivialization φ maps each fiber $\pi^{-1}(b)$ of π over a point $b \in U$ to the subset $\{b\} \times \pi^{-1}(\overline{b})$. Thus, it provides an identification of all the fibers of π over the open set U with the fiber over $\overline{b}$, and the set of all these fibers has a product structure. It follows that all fibers of π over U are diffeomorphic.

It also follows that, *if B is connected, then all the fibers of π are diffeomorphic to each other.* (Given two points $b, b' \in B$ there is a path that connects b and b', and any such path can be covered with a finite number of local trivializations). In such a case, we may choose one of the fibers of π (or a manifold diffeomorphic to it) as *typical fiber* F and, via composition with the diffeomorphisms between the fibers and F, replace the local trivializations with maps

$$\varphi : \pi^{-1}(U) \to U \times F$$

which satisfy $\varphi|_{\pi^{-1}(U)} = p_1 \circ \varphi$, and are also called local trivializations. We will usually do this way.

Examples: 1. Any cartesian product $B \times F$ with projection $\pi : B \times F \to B$, $(b, f) \mapsto b$. In this case there is a global trivialization $\varphi : \pi^{-1}(B) = B \times F \to B \times F$, which is the identity.

2. The tangent bundle of a manifold. The local trivializations are provided by the 'lifted coordinates' (Section A.1.A).

3. The Möbius strip, the Klein bottle, the Hopf fibration $\pi : \mathrm{SO}(3) \to S^2$, ... are examples of 'non trivial' locally trivial fibrations, which (globally) are not cartesian products.

4. Not all submersions $\pi : M \to B$ are locally trivial fibrations. For instance, the map $\pi(x,y) = x$ from $M = \mathbb{R}^2 \setminus \{(0,0)\} \ni (x,y)$ to $B = \mathbb{R} \ni x$ is a submersion, but its fibers are all diffeomorphic to $\mathbb{R}$ except $\pi^{-1}(0)$, which is diffeomorphic to $\mathbb{R} \setminus \{0\}$, or $\mathbb{R} \cup \mathbb{R}$. Thus (M, B, π) is not a locally trivial fibration.

As the last example shows, a difference between a surjective submersion and a locally trivial fibration is that in the former case the fibers need not all have the same structure. But there is more than that—particularly when the fibers are not compact.

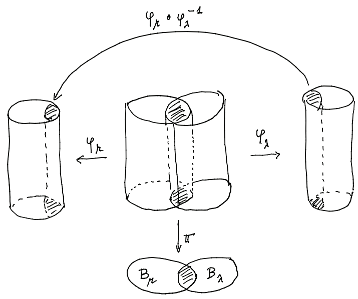
Informally, a locally trivial fibration has a product structure which is 'local in the base' but 'global in the fibers': (small) 'packets' $\pi^{-1}(U)$ of fibers, over (sufficiently small) open sets U in the base, can be 'straightened out' to the product $U \times F$. On the other hand, for a submersion, the submersion theorem guarantees that this can be done only locally in the base *and* in the fibers.

There are various sufficient conditions under which a submersion $\pi : M \to B$ is a locally trivial fibration (Ehresmann's fibration theorem and extensions): 1. π is a proper map (automatic, if M is compact). 2. The fibers of π are compact and all have the same number of connected components (in particular, they are connected). 3. The fibers of π are diffeomorphic to $\mathbb{R}^p$, $p \geq 1$.

Locally trivial fibrations may have additional structures. One way to define them is in terms of the transition functions between local trivializations.

Definition 2.2.7 *An* atlas by local trivializations *of a locally trivial fibration (M, B, π) with typical fiber F is a collection of local trivializations $\varphi_\lambda : \pi^{-1}(B_\lambda) \to B_\lambda \times F$, $\lambda \in \Lambda$ (some index set) whose domains cover M, namely*

$$B = \cup_\lambda B_\lambda \, .$$



The transition function *between two local trivializations $\varphi_\lambda : \pi^{-1}(B_\lambda) \to B_\lambda \times F$ and $\varphi_\mu : \pi^{-1}(B_\mu) \to B_\mu \times F$ whose domain have nonempty intersection ($B_\lambda \cap B_\mu \neq \emptyset$) is the map*

$$\varphi_\mu \circ \varphi_\lambda^{-1} \; : \; (B_\lambda \cap B_\mu) \times F \to (B_\lambda \cap B_\mu) \times F \, . \tag{2.2.5}$$

Each transition function is a diffeomorphism, and has the form

$$\varphi_\mu \circ \varphi_\lambda^{-1}(b, y) = (b, \mathcal{F}_{\mu\lambda}(b, y))$$

with a certain map $\mathcal{F}_{\mu\lambda} : (B_\lambda \cap B_\mu) \times F \to F$ which, for any fixed b, is a diffeomorphism $\mathcal{F}_{\mu\lambda}(b, \cdot) : F \to F$.

Definition 2.2.8 *A locally trivial fibration is said to be a G-*principal bundle *if*

i. its typical fiber is a Lie group G, and *ii.* *it has an atlas by local trivializations in which all maps $\mathcal{F}_{\mu\lambda}$ as in (2.2.5) act by right-translations on G: for each $\lambda, \mu \in \Lambda$ there exist a map*

$$s_{\mu\lambda} : B_\lambda \cap B_\mu \to G$$

such that $\mathcal{F}_{\mu\lambda}(b, g) = g s_{\mu\lambda}(b)$.

Example: A *vector bundle* is a particular case of principal bundle, with typical fiber a vector space on which the transition maps act linearly. The tangent and cotangent bundles of a manifold are examples of this situation.

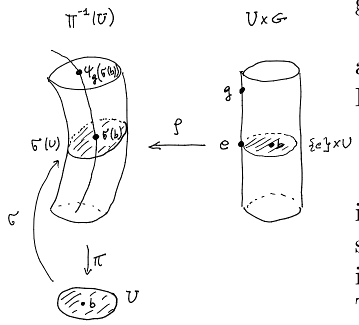
2.2.D Lie group actions and principal bundles. The orbits of free and proper Lie group actions are the fibers of principal bundles:

Proposition 2.2.9 *Consider a free and proper action Ψ of a Lie group G on a manifold M. Let $\pi : M \to M/G$ be the quotient map. Then, $(M, M/G, \pi)$ is a G-principal bundle.*

Proof. Let us write B for M/G. Choose a point $b^* \in B$, a point $m^* \in \pi^{-1}(b^*)$ and a local section

$$\sigma : U \to M\,, \qquad b \to \sigma(b)\,,$$

of $\pi : M \to B$ through m^*, namely a (smooth) map from a neighbourhood $U \subseteq B$ of b^* into M, such that $\pi \circ \sigma = \mathrm{id}_U$ and $m^* = \sigma(b^*)$. (Its existence is granted by the submersion theorem, see the Exercises).



For any $b \in B$, the orbit map $\Psi^{\sigma(b)} : G \to \pi^{-1}(b)$, which to each $g \in G$ associates the point $\Psi^{\sigma(b)}(g) = \Psi_g(\sigma(b))$ is a diffeomorphism (see the proof of Proposition 2.1.10). It follows that the map

$$\rho : U \times G \to \pi^{-1}(U)\,, \qquad (b, g) \mapsto \Psi_g(\sigma(b))\,,$$

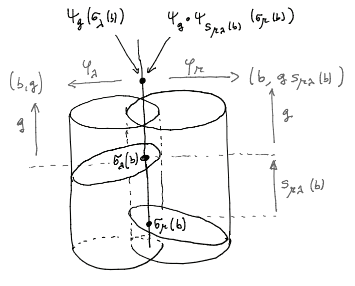
is bijective and, being a composition of smooth maps, is smooth. Its inverse is smooth as well, being given by $\rho^{-1}(m) = \big(\pi(m), (\Psi^{\sigma(\pi(m))})^{-1}(m)\big)$. Hence, ρ is a diffeomorphism. Moreover, ρ maps each set $\{b\} \times G$ onto the fiber $\pi^{-1}(b)$. Thus, its inverse $\varphi := \rho^{-1}$

$$\varphi : \pi^{-1}(U) \to U \times G$$

is a local trivialization for $\pi : M \to B$. By the arbitrariness of b^*, this shows that (M, B, π) is a locally trivial fibration. Note that

$$\varphi\big(\Psi_g(\sigma(b)\big) \;=\; (b, g) \qquad \forall b, g\,.$$

Consider two local trivializations $\varphi_\lambda : \pi^{-1}(B_\lambda) \to B_\lambda \times G$ and $\varphi_\mu : \pi^{-1}(B_\mu) \to B_\mu \times G$ with local sections $\sigma_\lambda : B_\lambda \to M$ and $\sigma_\mu : B_\mu \to M$. Assume $B_\lambda \cap B_\mu \neq \emptyset$. For any $b \in B_\lambda \cap B_\mu$, the points $\sigma_\lambda(b)$ and $\sigma_\mu(b)$



belong to the same orbit and—since the action is free—there exists a unique $s_{\mu\lambda}(b) \in G$ such that

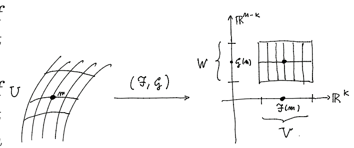
$$\sigma_\lambda(b) = \Psi_{s_{\mu\lambda}(b)}(\sigma_\mu(b)) .$$

This defines a map $s_{\mu\lambda} : B_\lambda \cap B_\mu \to G$ which is clearly smooth. Thus

$$\begin{aligned}\varphi_\mu \circ \varphi_\lambda^{-1}(b,g) &= \varphi_\mu\big(\Psi_g(\sigma_\lambda(b))\big) = \varphi_\mu\big(\Psi_g \circ \Psi_{s_{\mu\lambda}(b)}(\sigma_\mu(b))\big) \\ &= \varphi_\lambda\big(\Psi_{g s_{\mu\lambda}(b)}(\sigma_\mu(b))\big) = \big(b, g s_{\mu\lambda}(b)\big) .\end{aligned}$$

This proves that (M, B, π) is a G-principal bundle. ■

Exercises 2.2.3 *(i) Show that if $\pi : M \to B$ is a surjective submersion, then any point of M belongs to a local section of $\pi : M \to B$. [Hints: This is a local question and can be dealt with in coordinates. The submersion theorem ensures that if $\mathcal{F} = (\mathcal{F}_1, \dots, \mathcal{F}_k) : M \to \mathbb{R}^k$ is a submersion, then for any $m \in M$ there exist neighbourhoods U of m in M and V of $\mathcal{F}(m)$ in $\mathbb{R}^k$, an open set W in $\mathbb{R}^{n-k}$ and a map $\mathcal{G} = (\mathcal{G}_1, \dots, \mathcal{G}_{n-k}) : U \to W$ such that $(\mathcal{F}|_V, \mathcal{G}) : U \to V \times W$ is a diffeomorphism. Verify that $\sigma : V \to U$, $\sigma(x) = (x, \mathcal{G}(m))$, is a local section.]*



2.2.E The reconstruction equation. If $\pi : M \to M/G =: B$ is the quotient map associated to a free and proper action Ψ of a Lie group G on M, then the local trivializations of the G-principal bundle (M, B, π) provide 'semi-global' models for M, which are very useful to study the dynamics of Ψ.

First, we note how actions change under diffeomorphisms. If $\Psi : G \times M \to M$ is an action and $\varphi : M \to N$ is a diffeomorphism, then the map $\Psi^\varphi : G \times N \to N$ defined by

$$\Psi^\varphi_g := \varphi \circ \Psi_g \circ \varphi^{-1} \qquad \forall g \in G$$

$$\begin{tikzcd} M \arrow[r, "\varphi"] \arrow[d, "\Psi_g"'] & N \arrow[d, "\Psi^\varphi_g"] \\ M \arrow[r, "\varphi"] & N \end{tikzcd}$$

is an action on N, and φ is said to *intertwine* Ψ and Ψ^φ. If $X \in \mathfrak{X}(M)$ is Ψ-invariant then $\varphi_* X \in \mathfrak{X}(N)$ is Ψ^φ-invariant (see the exercises).

Example: A local trivialization $\varphi : \pi^{-1}(U) \to U \times G$ of the principal bundle (M, B, π) defined by a free and proper action $\Psi : G \times M \to M$ intertwines Ψ to the action Ψ^φ of G on $U \times G$ given by

$$\Psi^\varphi_g(b, h) = (b, L_g h) , \tag{2.2.6}$$

namely, the factor U is left fixed and G acts by left-translation on itself. Indeed, if φ is built with reference to a local section σ of (M, B, π), then $\Psi^\varphi_g(b,h) = \varphi \circ \Psi_g \circ \varphi^{-1}(b,h) = \varphi \circ \Psi_g \circ \Psi_h(\sigma(b)) = \varphi \circ \Psi_{gh}(\sigma(b)) = (b, gh)$.

Proposition 2.2.10 *Assume that $X \in \mathfrak{X}(M)$ is invariant under a free and proper action $\Psi : G \times M \to M$. Let $\varphi : \pi^{-1}(U) \to U \times G$ be a local trivialization of the principal bundle (M, B, π). Then, there exist a map*

$$\xi : U \to \mathfrak{g}$$

and a vector field $\overline{X} \in \mathfrak{X}(U)$ such that

$$(\varphi_* X)(b,g) = \big(\overline{X}(b)\,,\, T_e L_g \cdot \xi(b)\big) \qquad \forall\, b \in U,\, g \in G\,. \tag{2.2.7}$$

Proof. We may identify $T_{(b,g)}(U \times G)$ and $T_b U \times T_g G$. Accordingly, we write $\varphi_* X = (Y, Z)$ with maps $Y : U \times G \to TU$ and $Z : U \times G \to TG$ such that $Y(b,h) \in T_b B$ and $Z(b,h) \in T_h G$ for all b, h.

As seen above, $\varphi_* X$ is invariant under the action Ψ^φ of G on $U \times G$ given by (2.2.6). For any $y \in T_b U$ and $z \in T_g G$,

$$T_{(b,h)} \Psi^\varphi_g \cdot (y, z) = (y, T_h L_g \cdot z)$$

and

$$\begin{aligned} \varphi_* X \circ \Psi^\varphi_g(b,h) \;&=\; (\varphi_* X)(b, gh) \;=\; \big(Y(b,gh), Z(b,gh)\big) \\ T_{(b,h)} \Psi^\varphi_g \cdot \big(\varphi_* X(b,h)\big) \;&=\; \big(Y(b,h), T_h L_g \cdot Z(b,h)\big)\,. \end{aligned}$$

Thus, Y and Z satisfy, for all b, h, g,

$$Y(b, gh) = Y(b,h)\,, \qquad Z(b,gh) = T_h L_g \cdot Z(b,h)\,.$$

Since the action by left translations is transitive, the first equality shows that Y is independent of g. Thus Y can be regarded as a vector field $\overline{X} : U \to TU$. For $h = e$, the second equality gives $Z(b,g) = T_e L_g \cdot Z(b,e) = T_e L_g \cdot \xi(b)$ with

$$\xi(b) := Z(b,e) \in T_e G\,.$$

The proof is concluded by observing that there are no other conditions on $\overline{X}$ and ξ: for any choice of them, the vector field (2.2.7) is Ψ^φ-invariant. ■

Proposition 2.2.10 provides a semiglobal model for invariant dynamics. If a vector field X on a manifold M is invariant under a (free and proper) action of a Lie group G, then in a neighbourhood $\pi^{-1}(U)$ of any G-orbit the differential equation $\dot{m} = X(m)$ can be written as the system

$$\dot{b} = \overline{X}(b)\,, \qquad \dot{g} = T_e L_g \cdot \xi(b) \tag{2.2.8}$$

on $B \times G$, with a vector field $\overline{X}$ on $B = M/G$ and a map $\xi : U \to \mathfrak{g}$. The first equation is the *reduced equation* and the second is the *reconstruction equation.* The integral curves of the reduced equation describe how the integral curves of X move from G-orbit to G-orbit. The motion 'along' the G-orbits of each integral curve of X is described the reconstruction equation.

The (geometric) fact that the reduced equation is given by a vector field on the quotient space corresponds to the (analytic) fact that such an equation is decoupled from the reconstruction equation and implies that its integral curves can be determined without knowing those of the reconstruction equation.

This results in a 'reduction-reconstruction' strategy to integrate—or at least to study, given that differential equations usually cannot be integrated—a G-invariant differential equation. Given an initial datum $(b_0, g_0) \in U \times G$, the ensuing solution $t \mapsto (b(t), g(t))$ of system (2.2.8) can (in principle) be determined in two steps: (1) $t \mapsto b(t)$ is the solution of the reduced equation with initial datum $b(0) = b_0$. (2) $t \mapsto g(t)$ is the solution of the time-dependent equation

$$\dot{g} = T_e L_g \cdot \xi(b(t)) \tag{2.2.9}$$

on G with initial datum $g(0) = g_0$.

$$\begin{tikzcd}
(b_0, g_0) \arrow[r, dotted] \arrow[d] & t \mapsto (b(t), g(t)) \\
b_0 \arrow[r, "\text{Reduced equation}"] & t \mapsto b(t) \arrow[u, "\text{Reconstruction equation}"']
\end{tikzcd}$$

Equation (2.2.9) can also be written

$$\dot{g} = X_{\xi(b(t))}(g)$$

where X_ξ is the left-invariant vector field on G determined by the Lie algebra element ξ. Thus, along a solution of the reduced equation, the reconstruction equation is the simplest possible differential equation on a Lie group—that given by a left-invariant vector field, only time-dependent. In particular, it depends only on the group G, not on the particular G-invariant system under consideration. What depends on the details of the considered system are the reduced vector field and the map $\xi : U \to \mathfrak{g}$, while the properties of the motion along the orbit do to a large extent depend only on the group.

We warn that, in practice, due to its time-dependency, integrating the reconstruction equation for a generic reduced motion $t \mapsto b(t)$ may be prohibitive. However, it is the *qualitative properties* of the reconstructed motions that matter more, and in certain cases they may be determined.

Examples: We write the reconstruction equation for an abelian group and for the Lie subgroups of GL(n). We assume that the reduced motion $t \mapsto b(t)$ is known and write

$$\tilde{\xi}(t) := \xi(b(t)) \,.$$

1. $G = \mathbb{R}^n \ni x$. For any t, $\tilde{\xi}(t) \in \mathbb{R}^n$. The left-invariant vector fields are the constant vector fields and the reconstruction equation has the form

$$\dot{x} = \tilde{\xi}(t) \,.$$

Its integration is straightforward:

$$x(t) = x(0) + \int_0^t \tilde{\xi}(s) ds \,.$$

If $G = \mathbb{T}^n \ni \alpha\,(\mathrm{mod}1)$, then $\alpha(t) = \alpha(0) + \int_0^t \tilde{\xi}(s)ds$ (mod1).

2. Let now G be a Lie subgroup of GL(n). The Lie algebra $\mathfrak{g}$ is a subspace of $\mathrm{L}(n)$. Let us write $\xi(t) =: V(t) \in \mathfrak{g} \subseteq \mathrm{L}(n)$. Then $X_{V(t)}(A) = AV(t)$ for all $A \in G$ and the reconstruction equation is

$$\dot{A} = AV(t)\,, \qquad A \in G \subseteq \mathrm{GL(n)}\,.$$

This is a matrix linear time-dependent equation. Notwithstanding its linearity, the time-dependency can make it difficult or even impossible to solve it analytically. Specifically, its solutions are known in two cases:

i. $t \mapsto V(t) =: V$ is constant. Then, $A(t) = A(0)\exp(Vt)$.

ii. $t \mapsto V(t)$ is T-periodic ($T > 0$). Floquet theory ensures that there exists a time-dependent periodic (of period T or $2T$) linear isomorphism $t \mapsto S(t) \in$ GL(n) such that $A \mapsto B = S(t)A$ conjugates $\dot{A} = AV(t)$ to a constant coefficient linear differential equation $\dot{B} = BW$, $W \in \mathrm{L}(n)$.

These two cases correspond to the general cases of relative equilibria and relative periodic orbits.

Exercises 2.2.4 *(i) In elementary physics courses, the free fall equation $\ddot{z} = -g$, $z \in \mathbb{R}$ (and g a constant) is integrated with two consecutive integrations to obtain first $\dot{z}(t) = v_0 - gt$ and then $z(t) = z_0 + v_0 t - \frac{1}{2}gt^2$. Write this second order equation as a first order equation in $T\mathbb{R}$ and, exploiting an $\mathbb{R}$-invariance of it, interpret the above integration procedure as an instance of the reduction-reconstruction procedure.*

(ii) A map $\varphi : M \to N$ (not necessarily a diffeomorphism) is said to intertwine an action Ψ^M *of G on M and an action Ψ^N of G on N if*

$$\varphi \circ \Psi^M_g = \Psi^N_g \circ \varphi \qquad \forall g \in G\,.$$

Show that if $X^M \in \mathfrak{X}(M)$ is Ψ^M-invariant, $X^N \in \mathfrak{X}(N)$ is φ-related to X^M and φ is surjective, then X^N is Ψ^N-invariant. [Hints: you know that $\varphi \circ \Psi^M_g = \Psi^N_g \circ \varphi$, $T\varphi \circ X^M = X^N \circ \varphi$ and $T\Psi_g \circ X^M = X^M \circ \Psi_g$. Compute $[(\Psi^N_g)_* X^N] \circ \Psi^N_g \circ \varphi$ and show that it equals $X^N \circ \Psi^N_g \circ \varphi$. Then use the surjectivity of φ and of Ψ^N_g].

(iii) Show that, within the proof of Proposition 2.2.10, the fact that Y is independent of b could also be deduced from the fact that, by Proposition 2.2.2, X is π-related to a vector field on M/G.

2.3 The dynamics in relative equilibria

2.3.A Relative equilibria

Definition 2.3.1 *A* relative equilibrium *of a G-invariant vector field $X \in \mathfrak{X}(M)$ is a G-orbit that projects over an equilibrium of the reduced vector field.*

Relative equilibria are important for two reasons: they are often the starting point of the analysis of a system, and they shed some light on the integrability of the ODEs. The dynamics in relative equilibria of *compact* groups is known from the 1980's, thanks to the work of Krupa, Fields and others. The case of non-compact group is less completely understood.

We restrict to the case of a free and proper action $\Psi : G\times M \to M$. Assume that $X \in \mathfrak{X}(M)$ is Ψ-invariant and that the reduced vector field $\overline{X}$ vanishes at a point b^* of the reduced space $B = M/G$. Then the G-orbit $\pi^{-1}(b^*)$ is a relative equilibrium. We work in a local trivialization whose domain contains the relative equilibrium. The integral curves of X with initial data in $\pi^{-1}(b^*)$ belong to it and, in the local trivialization, are described by the reconstruction equation

$$\dot{g} = X_\xi(g) \tag{2.3.1}$$

with a (fixed, constant) vector $\xi \in \mathfrak{g}$ which depends on b^* and will be called the *generator of the dynamics* in the relative equilibrium. The flow of such a vector field is given by the exponential map of G,

$$\Phi_t^{X_\xi}(g) = g \, \exp_G(t\xi) \,, \qquad \forall \, t, g \,. \tag{2.3.2}$$

Thus, determining the dynamics in relative equilibria amounts to determine the properties of the group exponential. The consideration of $\mathbb{R}$ (where exp is the real exponential and $t \mapsto \exp(\xi t)$ goes to infinity) and of SO(2) or SO(3), where $t \mapsto \exp(\xi t)$ is trigonometric and is a rotation, shows that these properties depend on the group. The basic distinction is if G is compact or not. We will consider only the compact case.

Clearly, in order to determine the dynamics in a relative equilibrium we may consider only the integral curve of the identity, $g = e$ in (2.3.2), since the integral curves of others g are obtained by translations of this one.

Example: *Dynamics in relative equilibria of $G = \mathbb{T}^n$.* The exponential map of $\mathbb{T}^n \ni \langle\alpha\rangle := \alpha \,\mathrm{mod} 1$ is

$$\exp : \mathbb{R}^n \to \mathbb{T}^n \,, \qquad \xi \mapsto \exp(\xi) = \langle\xi\rangle \,.$$

We write $\omega = (\omega_1, \dots, \omega_n)$ for $\xi \in \mathbb{R}^n = \mathrm{lie}(\mathbb{T}^n)$ and consider the curve

$$t \mapsto \langle\omega t\rangle \,. \tag{2.3.3}$$

The topological properties of this curve are described by Krönecker theorem (that we have already met for $n = 2$ in section 1.1.D), according to which:

i. The image of the curve (2.3.3) is dense in $\mathbb{T}^n$ if and only if ω is 'nonresonant', namely $\nu \cdot \omega \neq 0$ for all $\nu \in \mathbb{Z}^n \setminus \{0\}$.
ii. If ω is 'resonant', namely $\omega \cdot \nu = 0$ for some nonzero integer vector ν, then the topological closure
$$\overline{\{\langle\omega t\rangle \,:\, t \in \mathbb{R}\}}$$
of the image of the curve (2.3.3) is a subtorus of $\mathbb{T}^n$ of a certain dimension $0 \leq p \leq n$.

The dimension p of the subtorus in the resonant case depends on the arithmetic properties of ω. If ω is resonant, then the set $R_\omega := \{\nu \in \mathbb{Z}^n : \nu \cdot \omega = 0\}$ is a nontrivial subgroup of $\mathbb{Z}^n$, hence a lattice of some rank $1 \le r \le n$,[5] and it is not difficult to show that $p = n - r$.

2.3.B Tori of compact Lie groups.

Definition 2.3.2 *i. A* torus *is a compact, connected and abelian Lie group of dimension* ≥ 1.

ii. A torus of a Lie group G *is a torus subgroup of* G.

Examples: 1. For any nonzero $\xi \in \mathbb{R}^3$, $\{\exp(t\hat{\xi}) : t \in \mathbb{R}\}$ is a one-dimensional torus of SO(3).

2. The abelian group $\mathbb{R}^n$, $n \ge 1$, does not possess any torus.

Proposition 2.3.3 *If G is compact and connected, then any point of G belongs to a torus of G.*

Proof. Fix $g \in G$, $g \ne e$. Since $\exp_G : \mathfrak{g} \to G$ is surjective, there exists a nonzero $\xi \in \mathfrak{g}$ such that $g = \exp\xi$. The one-parameter subgroup $\{\exp(t\xi) : t \in \mathbb{R}\}$ is abelian and (since exp is continuous) connected, but not necessarily closed. Its closure

$$T_\xi := \overline{\{\exp(t\xi) : t \in \mathbb{R}\}} \tag{2.3.4}$$

is an abelian subgroup as well. Indeed, if $h, k \in T_\xi$ there exist sequences $t_i, s_i \in \mathbb{R}$ such that $h = \lim_{i\to+\infty} \exp(t_i\xi)$ and $k = \lim_{i\to+\infty} \exp(s_i\xi)$. Therefore $hk = \big(\lim_i \exp(s_i\xi)\big)\big(\lim_i \exp(t_i\xi)\big) = \lim_i \big(\exp(t_i\xi)\exp(s_i\xi)\big) = \lim_i \exp((t_i+s_i)\xi)$ which implies that $hk \in T_\xi$ and that $hk = kh$. Moreover, T_ξ is connected because the closure of a connected set is connected. Hence, T_ξ is a torus and $g \in T_\xi$. Since $\xi \ne 0$, $\dim T_\xi > 0$. The proof is concluded by observing that e belongs to T_ξ. ■

Non-compact groups may or may not have tori. A given element of a compact group may belong to different tori. A *maximal torus* of a Lie group G is a torus of G which is not properly contained in any other torus.

Proposition 2.3.4 *All maximal tori of a compact and connected Lie group G are conjugate to each other. (If T and T' are two maximal tori of G, then $T' = gTg^{-1}$ for some $g \in G$).*

[5] A *lattice of rank* $r \ge 1$ *of* $\mathbb{R}^n$ is a subset of $\mathbb{R}^n$ formed by the linear combinations with integers coefficients of r linearly independent vectors of $\mathbb{R}^n$, which are said to be a *set of generators* of the lattice. The fact that any subgroup of $\mathbb{Z}^n$ is a lattice follows, for instance, from the elementary divisor theorem.

We do not prove this (non trivial, and not easy to prove) result. It makes the following definition meaningful:

Definition 2.3.5 *The* rank *of a compact and connected Lie group is the dimension of its maximal tori.*

For instance, SO(3) has rank one.

Proposition 2.3.6 *Every k-dimensional torus ($k \geq 1$) is isomorphic, as a Lie group, to $\mathbb{T}^k$.*

Proof. Let T be a torus of dimension $n \geq 1$. The exponential map $\exp_T : \mathrm{lie}(T) = \mathbb{R} \to T$ is surjective and, since T is abelian, is a Lie group homomorphism. This implies that its kernel

$$\Lambda := \ker(\exp_T) = \{\lambda \in \mathrm{lie}(T) : \exp_T(\lambda) = e\}$$

is a subgroup of $\mathrm{lie}(T) = \mathbb{R}^n$. Moreover, this also implies that $\exp_T$ is a local diffeomorphism (by Proposition 1.3.16 it is a local diffeomorphism at 0 and by Proposition 1.1.7 it has constant rank). In turn, this implies that Λ is a *discrete* subgroup of $\mathbb{R}^n$, namely, every point of Λ has a neighbourhood in which there are no other points of Λ. Indeed, for any $\lambda \in \Lambda$ there exist neighbourhoods $U \subset \mathbb{R}^n$ of λ and $V \subset T$ of e such that $\exp_T|_U : U \to V$ is a diffeomorphism; thus, since $\exp_T(\lambda) = e$, $U \cap \Lambda = \{\lambda\}$.

We use now the algebraic fact that any nontrivial discrete subgroup of $\mathbb{R}^n$ is a lattice of some rank r, $1 \leq r \leq n$ (see footnote 5 on page 58). Thus, there exist r linearly independent vectors $\lambda_1, \ldots, \lambda_r \in \mathbb{R}^n$ such that

$$\Lambda = \Big\{ \textstyle\sum_{j=1}^r \nu_j \lambda_j \in \mathbb{R}^n : \nu_1, \ldots, \nu_r \in \mathbb{Z} \Big\} .$$

Complete $\lambda_1, \ldots, \lambda_r$ to a basis $\lambda_1, \ldots, \lambda_n$ of $\mathbb{R}^n$. Then, there is a matrix $L \in$ GL(n) such that

$$\lambda_i = L e_i \, , \qquad i = 1, \ldots, n \, ,$$

where $e_1, \ldots, e_n$ are the vectors of the standard basis of $\mathbb{R}^n$. Consider now the group homomorphism

$$\Theta : \mathbb{T}^r \times \mathbb{R}^{n-r} \to T \, , \quad (\langle\alpha\rangle, y) \mapsto \exp_T \left(\textstyle\sum_{j=1}^r \alpha_j \lambda_j + \sum_{j=r+1}^n y_j \lambda_j \right) .$$

This map is well defined, namely, its value does not depend on the choice of $\alpha \in \mathbb{R}^r$ in the equivalence class $\langle\alpha\rangle = \alpha \bmod 1$. In fact, for any $\nu \in \mathbb{Z}^r$, since T is abelian,

$$\begin{aligned} &\exp_T \left(\textstyle\sum_{j=1}^r (\alpha_j + \nu_j) \lambda_j + \sum_{j=r+1}^n y_j \lambda_j \right) \\ &\quad = \exp_T \left(\textstyle\sum_{j=1}^r \alpha_j \lambda_j + \sum_{j=r+1}^n y_j \lambda_j \right) \exp_T \left(\textstyle\sum_{j=1}^r \nu_j \lambda_j \right) \\ &\quad = \exp_T \left(\textstyle\sum_{j=1}^r \alpha_j \lambda_j + \sum_{j=r+1}^n y_j \lambda_j \right) \end{aligned}$$

given that all $\nu_j \lambda_j \in \Lambda$. Moreover:

- Θ is surjective. Since $\exp_T$ is surjective and $\lambda_1, \dots, \lambda_n$ is a basis of $\mathrm{lie}(T)$, every $g \in T$ can be written as $\exp_T(\sum_j z_j\lambda_j)$ with $z_1, \dots, z_n \in \mathbb{R}$. And $\exp_T(\sum_j z_j\lambda_j) = \Theta(\langle z_1\rangle, \dots, \langle z_r\rangle, z_{r+1}, \dots, z_n)$.
- Θ is injective. If $\Theta(\langle\alpha\rangle, y) = \Theta(\langle\alpha'\rangle, y')$ then $\Theta(\langle\alpha - \alpha'\rangle, y - y') = e$ and $\sum_{j=1}^r (\alpha_j - \alpha_j')\lambda_j + \sum_{j=r+1}^n (y_j - y_j')\lambda_j \in \Lambda$. Thus $y = y'$ and $\alpha - \alpha' \in \mathbb{Z}^r$, hence $\langle\alpha\rangle = \langle\alpha'\rangle$.
- Θ is a local diffeomorphism. Fix a point $\langle\alpha\rangle \in \mathbb{T}^r$ and an open set U in $\mathbb{R}^r$ which is such that $\alpha \in \langle U\rangle := \{\langle x\rangle : x \in U\}$ and is sufficiently small so that the quotient map $\pi_U : U \to \langle U\rangle$, $x \mapsto \langle x\rangle$, is a diffeomorphism. Thus, $\varphi := \pi_U^{-1} \times \mathrm{id}_{\mathbb{R}^{n-r}} : \langle U\rangle \times \mathbb{R}^{n-r} \to U \times \mathbb{R}^{n-r}$ is a diffeomorphism (it is actually a coordinate system) and if we consider the linear isomorphism $j : \mathbb{R}^n \to \mathbb{R}^n = \mathrm{lie}(T)$, $j(z) = \sum_{j=1}^n z_j\lambda_j$, we can write $\Theta|_{\langle U\rangle\times\mathbb{R}^{n-r}} = \exp_T \circ j \circ \varphi$. Thus, restricted to a neighbourhood of each point of its domain, Θ is the composition of two diffeomorphisms and of a local diffeomorphism. This shows that it is a local diffeomorphism.

$$\begin{tikzcd} \langle U\rangle \times \mathbb{R}^{n-r} \arrow[r, "\Theta"] \arrow[d, "\pi_U^{-1}\times\mathrm{id}"'] & T \\ \mathbb{R}^n \arrow[r, "j"] & \mathbb{R}^n \arrow[u, "\exp_T"'] \end{tikzcd}$$

Being bijective and a local diffeomorphism, $\Theta : \mathbb{T}^r \times \mathbb{R}^{n-r} \to T$ is a diffeomorphism, and hence a Lie group isomorphism. Since T is compact, this implies that $r = n$ and $\Theta : \mathbb{T}^n \to T$. ■

Note that the proof of this Proposition shows that, if T is a torus, then

i. $\ker(\exp_T)$ is a lattice of $\mathbb{R}^n = \mathrm{lie}(T)$ of rank n.

ii. If $\lambda_1, \dots, \lambda_n$ is a set of generators of $\ker(\exp_T)$, then

$$\Theta : \mathbb{T}^n \to T\,, \qquad \langle\alpha\rangle \mapsto \exp_T\left(\textstyle\sum_{j=1}^n \alpha_j\lambda_j\right) \tag{2.3.5}$$

is a Lie group isomorphism.

If T is a torus of a Lie group G, then this remains true with $\exp_T$ replaced by $\exp_G$, given that the restriction of $\exp_G$ to T equals $\exp_T$ (Corollary 1.3.19).

2.3.C The dynamics in relative equilibria of compact groups. We may describe now the dynamics in a relative equilibrium of a G-invariant vector field $X \in \mathfrak{X}(M)$, when G is compact and connected. We identify the relative equilibrium with G via a local trivialization. The restriction of X to the relative equilibrium is conjugate to a left-invariant vector field $X_\xi \in \mathfrak{X}(G)$ on G, for some $\xi \in \mathfrak{g}$, with flow given by (2.3.2).

Proposition 2.3.7 *Assume that G is compact and connected. Fix $\xi \in \mathfrak{g}$ and let T_ξ be the torus (2.3.4) generated by ξ. Then, for any $g \in G$:*

i. gT_ξ is invariant under the flow of X_ξ.

ii. The restriction of the flow of X_ξ to gT_ξ is conjugate to a linear flow

$$(t, \langle\alpha\rangle) \mapsto \langle\alpha + \omega t\rangle$$

on $\mathbb{T}^k$, $k = \dim T_\xi$, with a frequency vector $\omega \in \mathbb{R}^k$ that depends only on ξ (not on $g \in G$).

Proof. (i) $\Phi_t^{X_\xi}(g) = g \exp_G(t\xi) = g \exp_{T_\xi}(t\xi) \in gT_\xi$ for all $t \in \mathbb{R}$.

(ii) Consider a set of generators $\lambda_1, \ldots, \lambda_k$ of $\ker(\exp_{T_\xi})$ and the associated Lie group isomorphism $\Theta : \mathbb{T}^k \to T_\xi$ as in (2.3.5). Write

$$\xi = \textstyle\sum_{j=1}^k \omega_j \lambda_j$$

with $\omega = (\omega_1, \ldots, \omega_k) \in \mathbb{R}^k$. For any $g \in T_\xi$, $g = \exp_T(\sum_j \alpha_j \lambda_j)$ and thus, given that T_ξ is abelian,

$$\Theta\big(\langle \alpha + t\omega \rangle\big) \;=\; \exp\big(\textstyle\sum_j (\alpha_j + t\omega_j)\lambda_j\big) \;=\; \exp(\alpha)\exp(t\xi) \;=\; g\exp(t\xi)\,.$$

Consider now a T_ξ-orbit $g_0 T_\xi$ different from T_ξ. The map $\Theta_{g_0} := L_{g_0} \circ \Theta : \mathbb{T}^k \to g_0 T_\xi$ is a diffeomorphism. If $g \in g_0 T_\xi$ then $g_0^{-1} g \in T_\xi$ and so $g_0^{-1} g = \Theta(\langle \alpha \rangle)$ for some $\langle \alpha \rangle \in \mathbb{T}^k$. Thus, $\Theta_{g_0}(\langle \alpha + \omega t \rangle) = L_{g_0}(g_0^{-1} g \exp(t\xi)) = g \exp(t\xi))$. ■

The torus T_ξ, its dimension k, and the frequencies $\omega \in \mathbb{R}^k$ are determined by the generator $\xi \in \mathfrak{g}$, and the motion in the torus is either nonresonant or resonant depending on ξ. Of course, $\dim T_\xi \le \operatorname{rank} T$.

Thus, in a relative equilibrium of a compact group, motions take place on submanifolds diffeomorphic to tori, whose dimension does not exceed the rank of the group. These submanifolds are called 'invariant tori' of X and give a partition of the relative equilibrium into disjoint submanifolds (each point of the relative equilibrium belongs to one of them and any two of them either coincide or are disjoint). This partition into invariant submanifolds is in fact a T_ξ-principal fibration, see next section.

Courses in Mechanics are a source of examples of relative equilibria (circular motion in the Kepler problem, horizontal motions of a spherical pendulum, the rotations of a vertically standing top, ...).

2.3.D Structure of relative equilibria of compact groups. If H is a subgroup of a group G, then the *left cosets* of H are the sets gH, $g \in G$, namely, the orbits of the (right) action by right-translations of H on G, which is a free action. Denote by $G/^r H$ the quotient space under this right action. It is well known from the algebra courses that if H is a normal subgroup then $G/^r H$ is a group, with the quotient map $p : G \to G/^r H$ being a group homomorphism, but this is not true otherwise. However, the structure which interests us is that, by Proposition 2.2.9,[6] if H is compact then $G/^r H$ is a manifold and $(G, G/^r H, p)$ is an H-principal bundle.

[6]which is valid for right actions as well

This implies that, if T is a k-dimensional torus of G, then

$$p : G \to G/^r T$$

is a T-principal bundle (and hence a $\mathbb{T}^k$-principal bundle). In particular, G is foliated by the submanifolds gT, $g \in G$, which are diffeomorphic to T. Thus, Proposition 2.3.7 implies that:

Proposition 2.3.8 *Assume that $X \in \mathfrak{X}(M)$ is invariant under a free action of a compact and connected Lie group G. Then, for every relative equilibrium $R = \pi^{-1}(b^*)$ of X, there exists a torus T of G and a locally trivial fibration*

$$p : R \to G/^r T$$

whose fibers are invariant and diffeomorphic to T, and the restriction of the flow of X to each of them is conjugate to a linear flow on $\mathbb{T}^{\dim T}$ whose frequencies depend only on T.

Proof. If $\xi \in \mathfrak{g}$ is the generator of the dynamics in the relative equilibrium, this follows from Proposition 2.3.7 with $T = T_\xi$ (or, in fact, any other larger torus of G that contains T_ξ). ■

Remarks: (i) Proposition 2.3.8 implies that the 'invariant tori' gT in a relative equilibrium R are the fibers of a submersion. If we coordinatize an open set U of the base $G/^r T$ we obtain $\dim G - \dim T$ functions $f_i : p^{-1}(U) \to \mathbb{R}$ which are independent and of whom the invariant tori in $p^{-1}(U)$ are the level sets. Thus, these functions are first integrals of the restriction of X to the relative equilibrium. In this situation, therefore, *a group action produces first integrals* (a number $\dim G - \dim T \geq \dim G - \operatorname{rank} G$ of them).

(ii) Proposition 2.3.8 follows from Proposition 2.3.7 with the torus $T = T_\xi$, with ξ a generator of the dynamics in the relative equilibrium. However, if T_ξ is not a maximal torus, then such a Proposition is true also with any torus T that contains T_ξ, in particular with a maximal torus that contains T_ξ. Replacing T_ξ with a larger torus T means, geometrically, to put together the fibers of $p : G \to G/^r T_\xi$ to form a fibration with larger fibers (and smaller base). Dynamically, of course, a number of frequencies on the larger torus will be zero. Thus, this possibility is dynamically unjustified; however, it will be convenient in the integrability scenario of next section.

(iii) A subtler question is if, in case ξ is resonant, it is possible to replace T_ξ with a smaller torus of G, so as to obtain a dense flow. The answer is essentially (but not quite: it might be necessary to pass to a cover) affirmative.

2.4 Integrability

2.4.A The role of first integrals. Integrability of a dynamical system may have different signatures and characterizations, both for its dynamical features and the mechanisms that lead to it. The situation is particularly well understood in classical mechanics, where, from a dynamical point of view, integrability is identified with the fact that all motions are quasi-periodic, and the origin of this situation is identified in the presence of symmetries.

The discussion of the previous section suggests a possible integrability 'scenario': *a G-invariant vector field $X \in \mathfrak{X}(M)$ whose reduced vector field $\overline{X}$ is zero has all its integral curves belonging to relative equilibria and thus, if G is compact, quasi-periodic* (= conjugate to linear motions on tori). As simple as it may be, this is not far from the most general situation. However, we are so far missing any understanding of mechanisms that force the reduced vector field to vanish. This is the existence of first integrals.

Let us assume, as usual, that a connected Lie group G acts freely and properly on a manifold M. Then, $B = M/G$ has dimension $p = \dim M - \dim G$ and, if $\overline{X}$ has p independent first integrals, then $\overline{X} = 0$ (see the exercises). The question is, where do first integrals of the reduced system come from?

First, we note that they are the same as *G-invariant* first integrals of the unreduced vector field on M. Recall from Definition 2.1.5 that a function $f : M \to \mathbb{R}$ is G-invariant if $f \circ \Psi_g = f$ for all $g \in G$.

Proposition 2.4.1 *Assume that a connected Lie group G acts freely and properly on a manifold M, with quotient map $\pi : M \to M/G$. Then:*

i. If a function $f : M \to \mathbb{R}$ is G-invariant then there exists a (smooth) function $\overline{f} : M/G \to \mathbb{R}$ such that $f = \overline{f} \circ \pi$.

ii. If $k \leq \dim M - \dim G$ functions $f_1, \ldots, f_k : M \to \mathbb{R}$ are G-invariant and independent, then the functions $\overline{f}_1, \ldots, \overline{f}_k : M/G \to \mathbb{R}$ such that $f_i = \overline{f}_i \circ \pi$ are independent.

iii. If a G-invariant function $f : M \to \mathbb{R}$ is a first integral of a G-invariant vector field X on M then $\overline{f}$ is a first integral of the reduced vector field $\overline{X}$.

Proof. (i) The existence of $\overline{f}$ follows from the fact that, being G-invariant, f is constant on the fibers of π. Concerning smoothness, consider a point $m \in M$. Since the quotient map π is a submersion, there are coordinates $(x, y) : U \to \mathbb{R}^p \times \mathbb{R}^{n-p}$ ($n = \dim M$, $p = \dim M/G$) in a neighbourhood U of m and coordinates $x : \pi(U) \to \mathbb{R}^p$ in the neighbourhood $\pi(U)$ of $\pi(m)$ such that $\pi^{\rm loc}(x, y) = x$. The representative $f^{\rm loc}$ of f is independent of the y coordinates. Thus $\overline{f}^{\rm loc} = f^{\rm loc}$, and the smoothness of $\overline{f}^{\rm loc}$ follows from that of $f^{\rm loc}$.

(ii) In local coordinates as in (i), the representatives of $\overline{f}_1, \ldots, \overline{f}_k$ equal those of $f_1, \ldots, f_k$.

(iii) $\overline{f} \circ \Phi_t^{\overline{X}} \circ \pi = \overline{f} \circ \pi \circ \Phi_t^X = f \circ \Phi_t^X = f = \overline{f} \circ \pi$ for all t. Since π is surjective, this proves that $\overline{f}$ is a first integral of $\overline{X}$. ■

Exercises 2.4.1 *(i) Show that if a vector field on a p-dimensional manifold has p everywhere functionally independent vector fields then it is identically zero. [Hints: Different ways. (1) The dimension of the level sets of a submersion ..... (2) $L_X = 0$ because ....]*

(ii) Prove a local converse of the previous statement: if $X \in \mathfrak{X}(M)$ is zero, then any point $m \in M$ has a neighbourhood V such that $X|_V$ has $\dim M$ independent first integrals. [Hint: use local coordinates].

2.4.B An integrability result. An immediate consequence of the analysis so far is that *if a vector field $X \in \mathfrak{X}(M)$ is invariant under a free action of a compact Lie group G and possesses $\dim M - \dim G$ independent G-invariant first integrals, then its flow is quasi-periodic.* More precisely, each G-orbit is fibered by X-invariant submanifolds diffeomorphic to tori, and the restriction of the flow of X to each of these submanifolds is conjugate to a linear flow on a torus.

However, the construction we have done so far of this structure applies to each G-orbit separately. For instance, the dimension of the invariant tori built in the proof of Proposition 2.3.8 depend on the G-orbit (the generator $\xi \in \mathfrak{g}$ depends on G, and the dimension of T_ξ is sensitive to changes in ξ). With a little bit of work it is possible to show that, as hinted in Remark ii in section 2.3.D, the construction of these tori can be done uniformly through the G-orbits, by taking all of them of dimension $= \operatorname{rank} G$. We limit ourselves to state this result in the simplest—and probably most important—case:

Proposition 2.4.2 *Assume that $\mathbb{T}^k$ acts freely on an n-dimensional manifold M. Assume that a vector field $X \in \mathfrak{X}(M)$*

i. Is $\mathbb{T}^k$-invariant

ii. Possesses $n-k$ independent $\mathbb{T}^k$-invariant first integrals.

Then, in any local trivialization $\varphi : \pi^{-1}(U) \to U \times \mathbb{T}^k \ni (b, \alpha)$ of the principal bundle $\pi : M \to M/\mathbb{T}^k$ induced by the action of $\mathbb{T}^k$, the flow of X is given by

$$\dot{b} = 0\,, \qquad \dot{\alpha} = \omega(b)$$

with a smooth map $\omega : U \to \mathbb{R}^k$.

Proof. This follows at once from (2.2.8). ■

Remark: In applications, it is rare that the group acts freely and the first integrals are independent in the entire phase space of the system; often, these conditions are satisfied in a typically 'large' (e.g., dense) open X-invariant submanifold M_* of M, and the statement applies to the restriction of X to M_*.

2.4.C An integrability criterion. Proposition 2.4.2 catches a good part of the current interpretation of integrability, but has the limit that, in applications, it is rare that a $\mathbb{T}^k$-action is manifest. Often, an $\mathbb{R}^k$-action (or a $\mathbb{T}^h \times \mathbb{R}^{k-h}$-action for some $0 \le h \le k$) is given, together some conditions on it and on the first integrals that imply that its orbits coincide with those of a $\mathbb{T}^k$-action. This situation is particularly frequent in Hamiltonian mechanics, where there are more peculiarities because even the $\mathbb{R}^k$-action is produced by the first integrals (see section 2.5.C), but is not restricted to the Hamiltonian setting. A clear statement of it outside the Hamiltonian setting seems to have been given for the first time around 1995 by O. Bogoyavlenskij. We give the standard statement of this result, where the symmetry is induced via the flows of commuting vector fields (see Example 9, section 2.1.B).

Definition 2.4.3 *A vector field Y is a* dynamical symmetry *of a vector field X if $[X,Y]=0$.*

Proposition 2.4.4 *Assume that a vector field X on a manifold M of dimension n has, for some $0<k<n$,*

(B1) $n-k$ independent first integrals $f_1,\dots,f_{n-k}$) ($L_X f_h = 0$ for all $h = 1,\dots,n-k$) whose common level sets are compact and connected.

(B2) k dynamical symmetries $Y_1,\dots,Y_k$ ($[Y_i,X]=0$ for all $i=1,\dots,k$) which are everywhere linearly independent, pairwise commuting,

$$[Y_i,Y_j]=0 \qquad \forall i,j=1,\dots,k\,,$$

and preserve the first integrals, namely

$$L_{Y_i} f_h = 0 \qquad \forall i=1,\dots,k\,, \quad \forall h=1,\dots,n-k\,.$$

Then:

i. Every level set of $f=(f_1,\dots,f_{n-k})$ is diffeomorphic to $\mathbb{T}^k$.

ii. The restriction of the flow of X to each level set of f is conjugate to a linear flow on $\mathbb{T}^k$.

Proof. (i) Fix a level set N of f. N is a compact and connected k-dimensional submanifold of M (because f is a submersion) and is invariant under the flow of X (because $f_1,\dots,f_k$ are first integrals of X). Moreover, the vector fields $Y_1,\dots,Y_k$ are tangent to N ($L_{Y_i} f_j = 0$ means that Y_i is tangent to the level sets of f_j) and, since N is compact, their restrictions to N are complete vector fields and define flows $\Phi^{Y_1},\dots,\Phi^{Y_k}$ on N. Given that $Y_1,\dots,Y_k$ pairwise commute, their flows pairwise commute and define the action

$$\Psi:\mathbb{R}^k\times N\to N\,, \qquad \Psi_\tau(m)=\Phi^{Y_1}_{\tau_1}\circ\cdots\circ\Phi^{Y_k}_{\tau_k}$$

of $\mathbb{R}^k$ on N.

We now show that, for each $m\in N$, the orbit map

$$\Psi^m:\mathbb{R}^k\to N\,, \qquad \Psi^m(\tau):=\Psi_\tau(m)\,,$$

is a local diffeomorphism and is surjective:

- The tangent map $T_0\Psi^m : \mathbb{R}^k \to T_mN$ satisfies

$$T_0\Psi^m \cdot e_1 = \frac{d}{dt}\big(\Phi_t^{Y_1} \circ \Phi_0^{Y_2} \cdots \circ \Phi_0^{Y_k}\big)\big|_{t=0} = Y_1(m)$$

  and, since the flows of $Y_1, \ldots, Y_k$ commute, $T_0\Psi^m \cdot e_i = Y_i(m)$ for all $i = 1, \ldots, k$. Thus, by the linear independence of $Y_1(m), \ldots, Y_k(m)$, $T_0\Psi^m$ has rank k and is an isomorphism. By the inverse function theorem, Ψ^m is a local diffeomorphism at $0 \in \mathbb{R}^k$. Being equivariant with respect to the (transitive) action of $\mathbb{R}^k$ on itself by translation and to the action Ψ on N (see the proof of Proposition 2.1.10), Ψ^m has constant rank. Thus $\Psi^m : \mathbb{R}^k \to N$ is a local diffeomorphism.
- Being a local diffeomorphism, Ψ^m is an open map (maps open sets into open sets) and so the orbits $\mathcal{O}_m = \Psi^m(\mathbb{R}^k)$, $m \in N$, are open subsets of N. Choose an $m \in N$ and assume, by contradiction, that Ψ^m is not surjective. Then, $N \setminus \mathcal{O}_m$ is nonempty. But $N \setminus \mathcal{O}_m$ is a union of Ψ-orbits, and hence is an open set (any union of open sets is open). Thus, N is the union of two disjoint open sets, $\mathcal{O}_m$ and $N \setminus \mathcal{O}_m$. But this is impossible if N is connected.

Note that the surjectivity of the orbit map means that the action Ψ is transitive.

Fix now a point $m \in N$. The orbit map Ψ^m is not injective (if it was, then, being a surjective local diffeomorphism it would be a diffeomorphism, which is impossible because N is compact and $\mathbb{R}^k$ is not). Consequently, the action Ψ is not free and the isotropy subgroup $\Lambda \subseteq \mathbb{R}^k$ of m is non trivial. Since Ψ^m is a local diffeomorphism, Λ is a discrete subgroup of $\mathbb{R}^k$ and thus a lattice of $\mathbb{R}^k$ of rank $1 \le r \le k$ (see the proof of Proposition 2.3.6[7]). Choose generators $\lambda_1, \ldots, \lambda_r$ of Λ and complete them to a basis $\lambda_1, \ldots, \lambda_k$ of $\mathbb{R}^k$.

Proceeding as in the proof of Proposition 2.3.6, it is easy to check that the map

$$\widehat{\Psi}^m : \mathbb{T}^r \times \mathbb{R}^{k-r} \to N\,, \quad \widehat{\Psi}^m(\langle\tau'\rangle, \tau'') = \Psi^m\big(\textstyle\sum_{j=1}^r \tau'_j\lambda_j + \sum_{j=r+1}^k \tau''_j\lambda_j\big)$$

is well defined and smooth, that it inherits from Ψ^m the properties of being surjective and a local diffeomorphism, and that in addition it is injective. Hence, it is a diffeomorphism. This implies that $k = r$ and N is diffeomorphic to $\mathbb{T}^k$.

(ii) Let now L be the matrix such that

$$L^T e_i = \lambda_i\,, \qquad i = 1, \ldots, k\,,$$

[7] In fact, the core of the arguments at the base of Proposition 2.3.6 and of the present one could be unified. An action $\Psi : G \times M \to M$ is said *locally free* if there is a neighbourhood U of e in G such that $\Psi_g(m) \neq m$ for all $g \in U$, $g \neq e$. Thus, one can prove the following: *Let Ψ be a transitive and locally free action of $\mathbb{R}^k$ on a compact, connected k-dimensional manifold N. Then, there exists a transitive and free action $\tilde{\Psi}$ of $\mathbb{T}^k$ on N such that $\Psi_\tau = \tilde{\Psi}_{\langle\tau\rangle}$ for all $\tau \in \mathbb{R}^k$.* In particular, N is diffeomorphic to $\mathbb{T}^k$.

and consider the map

$$\widehat{\Psi} : \mathbb{T}^k \times N \to N\,, \quad (\langle\tau\rangle, m) \mapsto \widehat{\Psi}_{\langle\tau\rangle}(m) := \Psi_{L\tau}(m)\,.$$

Since $L\tau = \sum_{j=1}^k \tau_j \lambda_j$ for all $\tau \in \mathbb{R}^k$, $\widehat{\Psi}_{\langle\tau\rangle}(m) = \widehat{\Psi}^m(\langle\tau\rangle)$ and $\widehat{\Psi}$ is well defined. Moreover, it is an action of $\mathbb{T}^k$ on N which is transitive and free (because its orbit map $\widehat{\Psi}^m$ is surjective and injective. The vector field X is $\widehat{\Psi}$-invariant because it is Ψ-invariant $((\widehat{\Psi}_{\langle\tau\rangle})_* X = (\Psi_{L\tau})_* X = X)$. The statement now follows from Proposition 2.4.2. ■

There is also a 'semi-global' version of this result, which describes the structure of X not just on a single level set of the first integrals, but in an open neighbourhood of one of them—hence in a 'packet' of level sets. This ensures, in particular, that the frequencies $\omega = (\omega_1, \ldots, \omega_n)$ change smoothly from level set to level set of f:

Proposition 2.4.5 *Under the hypotheses of Proposition 2.4.4, $(M, f(M), f)$ is a locally trivial fibration with typical fiber $\mathbb{T}^k$. In each local trivialization $\varphi : f^{-1}(U) \to U \times \mathbb{T}^k \ni (b, \alpha)$, the flow of X is given by the equations*

$$\dot{b} = 0\,, \qquad \dot{\alpha} = \omega(b)$$

with a smooth map $\omega : U \to \mathbb{R}^k$.

Remarks: 1. In examples, it often happens that not all of the phase space of an integrable systems is filled by invariant tori. In such situations, the set M of Propositions 2.4.4 and 2.4.5 is a (typically 'large') subset of the phase space. At the boundary of M, the foliation by the invariant tori may have various types of singularities (in particular, invariant tori of smaller dimensions).

2. The hypothesis of connectedness of the level sets of the first integrals is not essential: Propositions 2.4.4 and 2.4.5 apply with M replaced by a neighbourhood of each connected component (however, the global structure of the foliation by the invariant tori might be more complicated than a locally trivial fibration).

2.5 The Hamiltonian case

2.5.A Hamiltonian systems. First, we recall a few facts about Hamiltonian systems, some of which were already seen in Examples 6 and 7 of section 1.2.A. As in that section, we consider the simple case in which the phase space is an open set M of $\mathbb{R}^{2d}$. Denote the coordinates in $\mathbb{R}^{2d}$ as $(q, p) = (q_1, \ldots, q_d, p_1, \ldots, p_d)$. The *Hamiltonian vector field* of a function $f : M \to \mathbb{R}$ is the vector field on M

$$X_f := \mathbb{J}\nabla f \;\in\; \mathfrak{X}(\mathbb{R}^{2d}) \tag{2.5.1}$$

where

$$\mathbb{J} = \begin{pmatrix} \mathbb{O}_d & \mathbb{I}_d \\ -\mathbb{I}_d & \mathbb{O}_d \end{pmatrix}$$

with $\mathbb{O}_d$ and $\mathbb{I}_d$ are the $d \times d$ zero and unit matrices. The function f is called the *Hamiltonian* of X_f. The matrix J is the so called symplectic unit and satisfies $\mathbb{J}^{-1} = \mathbb{J}^T = -\mathbb{J}$ and $\mathbb{J}^2 = -\mathbb{I}$. In coordinates,

$$X_f = \sum_i \left(\frac{\partial f}{\partial p_i} \partial_{q_i} - \frac{\partial f}{\partial q_i} \partial_{p_i} \right) .$$

Geometrically, $\mathbb{R}^{2d}$ is equipped with the bilinear anti-symmetric map $\sigma_{\mathbb{J}} : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to \mathbb{R}$, $(u, v) \mapsto \sigma_{\mathbb{J}}(u, v) := u \cdot \mathbb{J}v$. Since $\det \mathbb{J} = +1$, this map is nondegenerate (in the sense that $\sigma_{\mathbb{J}}(u, v) = 0$ for all $u \in \mathbb{R}^{2d}$ implies $v = 0$). Any bilinear anti-symmetric nondegenerate form σ on $\mathbb{R}^{2d}$ is called a *symplectic form* and $(\mathbb{R}^{2d}, \sigma)$ is called a *symplectic vector space*. Linear algebra shows that, if σ is a symplectic form on $\mathbb{R}^{2d}$, then there is a linear change of coordinates on $\mathbb{R}^{2d}$ under which σ takes the form $\sigma_{\mathbb{J}}$. It is therefore not restrictive to assume that this is the case, and from now on we write σ for $\sigma_{\mathbb{J}}$.

A symplectic structure on $\mathbb{R}^{2d}$ allows to associate vector fields to functions defined on (open subsets) of $\mathbb{R}^{2d}$. This is a specificity of the symplectic world and, as we now discuss, produces a link between first integrals and symmetries.

Remark: $(M \subseteq \mathbb{R}^{2d}, \sigma_{\mathbb{J}})$ is a special case of a symplectic manifold. A *symplectic form* on a $2d$-dimensional manifold M is a closed and nondegenerate differential 2-form σ on M. The Hamiltonian vector field of a function $h : M \to \mathbb{R}$ is the vector field X_h^σ defined by $\sigma(X_h^\sigma, \cdot) = -dh$. Darboux theorem ensures that, locally in M, there exist charts with coordinates $(q, p) = (q_1, \ldots, q_d, p_1, \ldots, p_d)$ such that the local representative of σ is $\sigma^{\mathrm{loc}} = \sum_{i=1}^d dp_i \wedge dq_i$. In Darboux coordinates, the matrix of the bilinear form σ^{loc} is $\mathbb{J}$ and the representative of X_h^σ is $\mathbb{J}\nabla h^{\mathrm{loc}}$.

2.5.B The algebras of first integrals and of Hamiltonian dynamical symmetries. The symplectic structure σ gives the set of functions on $\mathbb{R}^{2d}$ a Lie algebra structure. The *Poisson bracket* induced by σ is the bilinear map $\{\ ,\ \} : C^\infty(\mathbb{R}^{2d}) \times C^\infty(\mathbb{R}^{2d}) \to C^\infty(\mathbb{R}^{2d})$ defined by

$$\{f, g\} := -L_{X_f} g = \nabla f \cdot \mathbb{J} \nabla g \,. \tag{2.5.2}$$

Its coordinate expression was given in section 1.2.A. As there seen, the Poisson bracket satisfies the Jacobi identity and $(C^\infty(\mathbb{R}^{2d}), \{\ ,\ \})$ is a Lie algebra.

By (2.5.2), a function f is a first integral of a Hamiltonian vector field X_h if and only if the Poisson bracket $\{f, h\}$ is zero. In particular, every Hamiltonian vector field X_h has its own Hamiltonian h as a first integral.

A consequence of the Jacobi identity is that the Poisson bracket of two first integrals of a given Hamiltonian vector field is still a first integrals of that vector field: if $\{f,h\} = \{g,h\} = 0$ then

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

The interest of this fact is not that much that it allows producing new first integrals from known ones (because the new ones are likely not independent of the latter), but in that it shows that *the set of all first integrals of a given Hamiltonian vector field is a Lie subalgebra of* $(C^\infty(\mathbb{R}^{2d}), \{\ ,\ \})$.

Let now $\mathfrak{X}_H(\mathbb{R}^{2d})$ be the set of all Hamiltonian vector fields, which is a subspace of $\mathfrak{X}(\mathbb{R}^{2d})$.

Proposition 2.5.1 *For any two functions* $f_1, f_2 \in C^\infty(\mathbb{R}^{2d})$,

$$[X_{f_1}, X_{f_2}] = -X_{\{f_1,f_2\}}\,. \tag{2.5.3}$$

Proof. We use the fact that two vector fields are equal if they are equal as Lie derivative of functions (Proposition A.2.2). For any function f_3: $L_{[X_{f_1},X_{f_2}]}f_3 = L_{X_{f_1}}L_{X_{f_2}}f_3 - L_{X_{f_2}}L_{X_{f_1}}f_3 = -L_{X_{f_1}}\{f_2,f_3\} + L_{X_{f_2}}\{f_1,f_3\} = \{f_1,\{f_2,f_3\}\} + \{f_2,\{f_3,f_1\}\} = -\{f_3,\{f_1,f_2\}\} = -L_{X_{\{f_1,f_2\}}}f_3$. ∎

This has several consequences:

i. The set of all Hamiltonian vector fields is a Lie subalgebra of $(\mathfrak{X}(\mathbb{R}^{2d}), [\ ,\ ])$.

ii. The map $f \mapsto X_f$ is a Lie algebra anti-homomorphism from $(C^\infty(M), \{\ ,\ \})$ to $(\mathfrak{X}_H(\mathbb{R}^{2d}), [\ ,\ ])$.

And moreover:

i. If a function f is a first integral of a Hamiltonian vector field X_h ($\{f,h\} = 0$), then X_f is a dynamical symmetry of X_h ($[X_f, X_h] = 0$).

ii. The sets of all first integrals and of all Hamiltonian dynamical symmetries of a given Hamiltonian vector field X_h are anti-homomorphic Lie subalgebras of $(C^\infty(\mathbb{R}^{2d}), \{\ ,\ \})$ and of $(\mathfrak{X}_H(\mathbb{R}^{2d}), [\ ,\ ])$, respectively.

2.5.C The Liouville-Arnold theorem. At first sight, the characterization of integrability in the Hamiltonian case differs from that of generic systems because it is done in terms of first integrals alone, while dynamical symmetries are (apparently) absent. This is in fact due to the link between the two sets of objects in the symplectic setting. We illustrate this situation on the simplest Hamiltonian integrability result, which is known as the Liouville-Arnold theorem. Historically, this theorem predates—and is a predecessor of—Propositions 2.4.4 and 2.4.5.

Definition 2.5.2 *Two functions* $f, g \in C^\infty(\mathbb{R}^{2d})$ *are* in involution *if* $\{f,g\} = 0$.

By Proposition 2.5.1, the Hamiltonian vector fields of two functions in involution commute. Said differently, two functions are in involution if each one is a first integral of the Hamiltonian vector field of the other, and thus the Hamiltonian vector field of each of them is a dynamical symmetry of the Hamiltonian vector field of the other.

Proposition 2.5.3 (Liouville-Arnold theorem) *Assume that a Hamiltonian vector field X_h on an open set $M \subseteq \mathbb{R}^{2d}$ possesses d first integrals $f_1, \dots, f_d$ which are everywhere independent in M and pairwise in involution. Assume, moreover, that the map $f = (f_1, \dots, f_d)$ has compact and connected fibers. Then:*

i. The fibers of f are diffeomorphic to $\mathbb{T}^d$ and the restriction of the flow of X_h to each of them is conjugate to a linear flow on $\mathbb{T}^d$.

ii. $(M, f(M), f)$ is a locally trivial fibration with typical fiber $\mathbb{T}^d$, and for each $b \in f(M)$ there exist a neighbourhood U and a diffeomorphism $\psi : f^{-1}(U) \to V \times \mathbb{T}^d \ni (a, \alpha)$, with $V \subseteq \mathbb{R}^d$, such that

$$\psi^* \sigma = \sum_{i=1}^{d} da_i \wedge d\alpha_i$$

and $\psi^* f = a$ (namely, $f|_{f^{-1}(U)} = a \circ \psi$).

Our interest is to interpret statement i. in the light of Proposition 2.4.4. In the stated hypotheses, the Hamiltonian vector fields $X_{f_1}, \dots, X_{f_d}$:

- Are everywhere linearly independent, because $\nabla f_1, \dots, \nabla f_d$ are everywhere linearly independent (in view of the independence of $f_1, \dots, f_d$) and $X_{f_i} = \mathbb{J} \nabla f_i$ with an invertible matrix $\mathbb{J}$.
- Pairwise commute, because $f_1, \dots, f_d$ are pairwise in involution.
- Preserve the first integrals $f_1, \dots, f_d$, because $L_{X_{f_i}} f_j = \{f_i, f_j\} = 0$ for all i, j.

Statement i. thus follows from Proposition 2.4.4 with $n = 2d$, $k = d$, the $n - k = d$ first integrals $f_1, \dots, f_d$, and the $k = d$ dynamical symmetries $Y_1 = X_{f_1}, \dots, Y_d = X_{f_d}$. Thanks to the symplectic structure, and under the hypothesis of involutivity, the first integrals produce the commuting symmetries needed to produce the integrability.

We add a few comments:

First, seen from this point of view, it might appear that symmetry does not play any role in the integrability of Hamiltonian systems, which might seem to be entirely due to the existence of first integrals. That this is not the case is the content of the next chapter, where we will see that, in the Hamiltonian world (and more generally, in the variational world), it is symmetry that produces first integrals.

Second, it might appear that in a Hamiltonian system with d "degrees of freedom" (namely, with phase space of dimension $2d$), integrability always manifests itself through the existence of d integrals in involution and a fibration of invariant tori of dimension d. This is not the case, however, because a d-degrees of freedom Hamiltonian system may have more than d first integrals, and hence invariant tori of dimension $< d$. In the simplest case, assume that, in the situation of Proposition 2.5.3, there are additional first integrals $g = (g_1, \dots, g_{d-k})$ for some $0 < k < d$, such that $(f, g) : \mathbb{R}^{2d} \to \mathbb{R}^{2d-k}$ is a submersion. The intersections of the d-dimensional tori $f =$ const with the level sets of g are k-dimensional closed (hence compact) subsets of the d-dimensional invariant tori $f =$ const. Therefore, their connected components are diffeomorphic to $\mathbb{T}^k$. There are extensions of the Liouville-Arnold theorem ("non-commutative integrability") that generalize this situation to less obvious ones, and allow for invariant tori of any dimension n between $n = 1$ (periodic flow) and $n = d$.

Third, we stress again that the involutivity of the first integrals (which codifies the commutativity of the associated dynamical symmetries) plays a key role. As such, it has a name, even though used with slightly different declinations: A Hamiltonian system with d degrees of freedom is said to be *completely integrable* if it has d first integrals in involution which are independent in a significantly large (typically, either open and dense or of full measure) subset of the phase space. The Liouville-Arnold theorem describes the structure of completely integrable Hamiltonian systems under the additional hypothesis of compactness (and, but less important, connectedness) of the level sets of the first integrals. (The non-compact case is more complicated and has non quasi-periodic dynamics).

Examples: 1. Any Hamiltonian system with $d = 2$ degrees of freedom which has one first integral independent of the Hamiltonian is completely integrable. An example is the Hamiltonian

$$h(q,p) = \frac{1}{2}|p|^2 - V(|q|)\,, \qquad (q,p) \in \mathbb{R}^2 \times \mathbb{R}^2 \tag{2.5.4}$$

which describes a point particle in a plane subject to a central force field. The additional first integral is the angular momentum $L = q_1p_2 - q_2p_1$. Independence and compactness of (h, L) depend on the properties of V. If they are satisfied, then motions are quasi-periodic on two-dimensional tori. There are however two special cases in which there is a third independent first integral and all motions are periodic: these are $V(|q| = \frac{k}{2}|q|^2$ ($k > 0$, the harmonic oscillator) and $V(|q|) = -k|q|^{-1}$ ($k > 0$, Kepler system).

2. A point in a central force field in three-dimensional space, namely Hamiltonian (2.5.4) but in $\mathbb{R}^3 \times \mathbb{R}^3 \ni (q,p)$. There are four first integrals: h and the three components (K_1, K_2, K_3) of the angular momentum vector $K = q \times p$. K_1, K_2 and K_3 are not pairwise in involution, given that $\{K_1, K_2\} = K_3$ etc, but $|K|^2$ is in involution with each of them. Thus, the three functions $h, |K|^2, K_1$ (or $h, |K|^2, K_2$ or $h, |K|^2, K_3$) are in involution. Under the usual conditions

of independence and compactness, the Liouville-Arnold theorem describes in this case a fibration by invariant tori of dimension three, but the existence of four integrals implies that in fact there is a (finer) fibration by two-dimensional invariant tori, and motions are linear on them. Here too, for the harmonic and Keplerian potentials there is a fifth independent first integral and the dynamics is periodic.

Remark: The semi-global statement ii. is an analogue of Proposition 2.4.5, which however gives information on the symplectic structure that the fibration by the 'invariant tori' $f = \text{const}$ has in this case. The maps (a, α) are called *action-angle* variables and play an important role in Hamiltonian mechanics (in the study of small perturbations of integrable systems). The 'actions' $a = (a_1, \ldots, a_d)$ are a (semi-global) reparametrization of the first integrals (and the map ψ is a reparametrization of a local trivialization of the fibration $f : M \to f(M)$). It is easy to see that the local representative $h^{\text{loc}} = h \circ \psi^{-1}$ of the Hamiltonian h is independent of the angles α and X_h becomes $\dot{a} = 0$, $\dot{\alpha} = \frac{\partial h}{\partial a}(a)$. The expression of σ in the action-angle variables implies that the restriction of σ to the sets $a = \text{const}$, namely to the invariant tori, vanishes: this means that the invariant tori are 'Lagrangian submanifolds'.

(Version 2020.1)

Chapter 3

Symmetry and conservation laws

In classical mechanics, and in theoretical physics, symmetries play a role that goes beyond the one they have for generic differential equations: besides allowing reduction, they produce first integrals ("Noether theorem"). This is a very special situation, which is due to the *variational* nature of the equations of motion (Lagrange, or Hamilton, equations). A full comprehension of this topic would require the Hamiltonian formulation, but here we will limit ourselves to a first look in the Lagrangian context.

3.1 Variational systems and Noether theorem

3.1.A The variational principle and the Euler-Lagrange equations Let $Q \subset \mathbb{R}^d$ be an open set. Fix two points $q_0, q_1 \in Q$ and two 'times' $t_0 < t_1 \in \mathbb{R}$. Consider the set of all (smooth) parametrized curves

$$\Gamma := \big\{\gamma : [t_0, t_1] \to Q : \gamma(t_0) = q_0 , \ \gamma(t_1) = q_1\big\} .$$

This is an infinite dimensional affine space, with associated vector space

$$\Gamma_0 := \big\{\eta : [t_0, t_1] \to Q : \eta(t_0) = \eta(t_1) = 0\big\} .$$

Indeed, if $\gamma_1, \gamma_2 \in \Gamma$ then $\gamma_1 - \gamma_2 \in \Gamma_0$.

Traditionally, any function $F : \Gamma \to \mathbb{R}$ is called a *functional* and its value on a curve $\gamma \in \Gamma$ is denoted $F[\gamma]$. For instance, the *action functional* $A_L : \Gamma \to \mathbb{R}$ associated to a function $L : TQ \to \mathbb{R}$ is defined as

$$A_L[\gamma] := \int_{t_0}^{t_1} L\big(\gamma(t), \dot\gamma(t)\big) dt .$$

Definition 3.1.1 *Let $F : \Gamma \to \mathbb{R}$ be a functional.*

(i) F is Gateaux differentiable at $\gamma \in \Gamma$ *if, for any $\eta \in \Gamma_0$, the composed function $\lambda \to F(\gamma + \lambda\eta)$, which is a map from a real interval containing* 0 *to $\mathbb{R}$, is differentiable at* 0.

(ii) If F is Gateaux differentiable at γ, then the linear map

$$D^G_\gamma F : \Gamma_0 \to \mathbb{R}\,, \qquad \eta \mapsto (D^G_\gamma F)(\eta) := \frac{d}{d\lambda} F[\gamma + \lambda\,\eta]\big|_{\lambda=0}\,,$$

is the Gateaux differential of F at γ.

(ii) F is Gateaux differentiable *if it is Gateaux differentiable at each $\gamma \in \Gamma$.*

Example: The action functional A_L associated to any smooth function $L : TQ \to \mathbb{R}$ is Gauteaux differentiable and, for all $\gamma \in \Gamma$ and $\eta \in \Gamma_0$,

$$(D^G_\gamma A_L)(\eta) = -\sum_{i=1}^n \int_{t_0}^{t_1} \eta_i(t) \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q_i}\big(\gamma(t), \dot\gamma(t), t\big)\right) - \frac{\partial L}{\partial q_i}\big(\gamma(t), \dot\gamma(t), t\big)\right] dt\,.$$

The proof is a computation.

Definition 3.1.2 *Let $F : \Gamma \to \mathbb{R}$ be a Gateaux differentiable functional. A curve $\gamma \in \Gamma$* stationarizes F *if $D^G_\gamma F = 0$.*

Note that $D^G_\gamma F = 0$ means $(D^G_\gamma F)(\eta) = 0$ for all $\eta \in \Gamma_0$. The following result is classical:

Proposition 3.1.3 (Hamilton's principle) *Consider a smooth function $L : TQ \to \mathbb{R}$. A curve $\gamma \in \Gamma$ stationarizes the action functional A_L if and only if it satisfies the equations*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q_i}\big(\gamma(t), \dot\gamma(t)\big)\right) - \frac{\partial L}{\partial q_i}\big(\gamma(t), \dot\gamma(t)\big) \;=\; 0\,, \qquad i = 1, \dots, n\,. \tag{3.1.1}$$

The proof can be found in textbooks in analysis and classical mechanics.

The function L is called the *Lagrangian* of the variational problem. Depending on the context, equations (3.1.1) are called either Euler-Lagrange equations or *Lagrange equations* for the Lagrangian L.

This type of variational problems, and hence the Euler-Lagrange equations, emerge in classical mechanics (where the Lagrangian is often the difference between the kinetic energy and the potential energy of the system), in differential geometry (geodesics stationarize the length functional), in innumerable areas of theoretical physics, in control theory for ODEs, in optimization, etc.

Remarks: (i) The Euler-Lagrange equations are second order equations on Q. Writing them down explicitly, one sees that they can be put in normal form ($\ddot{q} = ...$) if the Lagrangian is such that, for all $(q, \dot{q})$, the matrix $\frac{\partial^2 L}{\partial \dot{q} \partial \dot{q}}(q, \dot{q})$ is invertible. In such a case the Lagrangian is said to ben *regular* and the Euler-Lagrange equations can be written as a first order system ($\dot{q} = v$, $\dot{v} = ...$) on TQ, namely, as a vector field on TQ, that we will denote here X^L.

(ii) $D^G_\gamma F(\eta)$ is also called the *variation* of F at γ along η, hence the term 'calculus of variations'.

(iii) Formulations of the calculus of variations on manifolds are also possible.

(iv) A curve that stationarizes a functional might not be a minimum.

3.1.B Noether theorem. Consider now an action Ψ of a Lie group G on Q. Let $\xi^Q = \frac{d}{d\lambda}\Psi_{\exp(\lambda\xi)}|_{\lambda=0}$ be the infinitesimal generator of the action associated to a vector $\xi \in \mathfrak{g}$ (Definition 2.1.7) and Ψ^{TQ} be the tangent lift of Ψ to TQ (section 2.1.B, example 14). Recall that, for all $g \in G$,

$$\Psi^{TQ}_g = T\Psi_g$$

and $\Psi^{TQ}_g(q, \dot{q}) = (\Psi_g(q), \Psi'_g(q)\dot{q}))$ for all $(q, \dot{q}) \in TQ$.

Proposition 3.1.4 (Noether theorem) *If $L : TQ \to \mathbb{R}$ is Ψ^{TQ}-invariant, then for any $\xi \in \mathfrak{g}$ the function*

$$J^L_\xi := \frac{\partial L}{\partial \dot{q}} \xi^Q \tag{3.1.2}$$

($= \sum_{i=1}^d \frac{\partial L}{\partial \dot{q}_i}\xi^Q_i$) is a first integral of the Euler-Lagrange equations for the Lagrangian L.

Proof. Let $t \mapsto q_t$ be a solution of Lagrange equations for the Lagrangian L. Note that $\frac{d}{dt}\Psi_{\exp(\lambda\xi)}(q_t) = \Psi'_{\exp(\lambda\xi)}(q_t)\dot{q}_t$ for all $t, \lambda \in \mathbb{R}$. Thus, the invariance of L implies

$$L\Big(\Psi_{\exp(\lambda\xi)}(q_t), \frac{d}{dt}\Psi_{\exp(\lambda\xi)}(q_t)\Big) = L(q_t, \dot{q}_t) \qquad \forall t, \lambda$$

and so

$$0 = \frac{d}{d\lambda} L\Big(\Psi_{\exp(\lambda\xi)}(q_t), \frac{d}{dt}\Psi_{\exp(\lambda\xi)}(q_t)\Big) = \frac{\partial L}{\partial q}\frac{d}{d\lambda}\Psi_{\exp(\lambda\xi)} + \frac{\partial L}{\partial \dot{q}}\frac{d}{d\lambda}\frac{d}{dt}\Psi_{\exp(\lambda\xi)}$$

where, in the last expression, the derivates of L are evaluated in $\big(\Psi_{\exp(\lambda\xi)}(q_t), \Psi'_{\exp(\lambda\xi)}(q_t)\dot{q}_t\big)$ and those of $\Psi_{\exp(\lambda\xi)}$ in q_t. Switching the order of the derivatives with respect to t and λ, evaluating everything at $\lambda = 0$ and noticing that for $\lambda = 0$ it is $\Psi_{\exp(\lambda\xi)} = \mathrm{id}$ and $\Psi'_{\exp(\lambda\xi)} = \mathbb{I}$ gives

$$\frac{\partial L}{\partial q}(q_t, \dot{q}_t)\,\xi(q_t) + \frac{\partial L}{\partial \dot{q}}(q_t, \dot{q}_t)\,\frac{d}{dt}\xi(q_t) = 0\,.$$

Since $t \mapsto q_t$ satisfies the Euler-Lagrange equations, here $\frac{\partial L}{\partial q} = \frac{d}{dt}\frac{\partial L}{\partial \dot q}$ and $\frac{d}{dt}\big(\frac{\partial L}{\partial \dot q}(q_t, \dot q_t)\xi(q_t)\big) = 0$. ■

The first integral (3.1.2) is called the *(Noetherian) momentum* of the action of the one-parameter subgroup generated by ξ. Note that every one-parameter subgroup produces a conserved momentum.

We recall that, in Lagrangian mechanics, the function $p_i := \frac{\partial L}{\partial \dot q_i}$ is called the *conjugate momentum* to the coordinate q_i. The momenta $(p_1, \dots, p_d)$ are the components of the differential 1-form $\sum_i p_i dq_i$ on Q, and for each $q, \dot q$, $p(q,\dot q) = (p_1(q,\dot q), \dots, p_d(q,\dot q))$ should accordingly be thought of not as a tangent vector but as a tangent covector, namely an element of T_q^*Q. To stress this fact, we shall write $\langle p, X\rangle$ for the pairing $p \cdot X = \sum_i p_i X_i$ with a vector field X on Q. Thus, the Noetherian momentum J_ξ^L associated to a Lie algebra vector ξ is the function

$$J_\xi^L = \big\langle p, \xi^Q \big\rangle : TQ \to \mathbb{R}\,.$$

Examples: 1. The Lagrangian of a mechanical system has often the form

$$L(q,\dot q) = \tfrac{1}{2}\dot q \cdot A(q)\dot q - V(q)\,, \qquad q \in \mathbb{R}^d\,, \tag{3.1.3}$$

with a symmetric, positive definite matrix A which depends on q. Note that all these Lagrangians are regular (in the sense of Remark (i) at the end of section 3.1.A). The conjugate momentum is $p(q,\dot q) = A(q)\dot q$. Given an action Ψ on $Q = \mathbb{R}^d$,

$$(L \circ T\Psi_g)(q,\dot q) \;=\; \frac{1}{2}\Psi_g'(q)\dot q \cdot A(q)\Psi_g'(q)\dot q - V(\Psi_g(q))$$

and this equals $L(q,\dot q)$ for all $g, q, \dot q$ if and only if

$$V(\Psi_g(q)) = V(q) \qquad \forall g, q \tag{3.1.4}$$

and

$$A(\Psi_g(q)) = \Psi_g'(q)^T A(q)\Psi_g'(q) \qquad \forall g, q\,. \tag{3.1.5}$$

(These are clearly sufficient conditions. Their necessity is proven by observing that for $\dot q = 0$ condition $L \circ \Psi_g^{TQ} = L$ gives (3.1.5), and (3.1.4) then follows from the fact that the matrix $(\Psi_g')^T A \Psi_g' - A$ is symmetric.) Thus, the invariance of the Lagrangian L under Ψ^{TQ} is equivalent to the invariance of the kinetic energy $\frac{1}{2}\dot q \cdot A(q)\dot q$ under Ψ^{TQ} (equation (3.1.5)) and of the potential energy V under Ψ (equation (3.1.4)).

2. For an unconstrained particle in cartesian coordinates, the Lagrangian is of the form (3.1.3) with $Q = \mathbb{R}^3$ (or $Q = \mathbb{R}^3 \setminus \{0\}$ if the potential energy is not defined at the origin) and $A(q) = m\mathbb{I}$ with $m > 0$ the mass of the particle. The momentum covector is $p = m\dot q$. Consider the action Ψ of $\mathbb{R}^3$ on $Q = \mathbb{R}^3$ by translations: $\Psi_\xi(q) = q + \xi$, $\Psi_\xi' = \mathbb{I}$ and $\Psi_\xi^{TQ}(q,\dot q) = (q+\xi, \dot q)$ (we write ξ for the group element, given that the group and the algebra can be identified). The kinetic energy is invariant under Ψ, while V is invariant if and only it is constant (no force acts on the system). The infinitesimal generator ξ^Q associated to a

vector ξ is the constant vector field $\xi^Q(q) = \xi$. The Noetherian momentum associated to a vector $\xi \in \mathbb{R}^3$ is $J_\xi^L = p \cdot \xi$. Note that the arbitrariness of ξ implies the constancy of the (co)vector $p = m\dot{q} \in \mathbb{R}^3$.

3. Same Lagrangian as in 2., but with the linear action of the rotation group SO(3) on $Q = \mathbb{R}^3$: $\Psi_R(q) = Rq$, $\Psi^{TQ}(q, \dot{q}) = (Rq, R\dot{q})$. The kinetic energy is invariant under rotations, while the potential energy V is invariant if and only if it is a function of the distance from the origin, namely

$$V(q) = \tilde{V}(|q|)$$

with a function $\tilde{V} : \mathbb{R} \to \mathbb{R}$. If we identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$, then the one-parameter subgroup associated to a Lie algebra vector $\xi \in \mathbb{R}^3$ is $\lambda \mapsto \Psi_{\exp(\lambda\widehat{\xi})}$ and the associated infinitesimal generator is

$$\xi^Q(q) \;=\; \frac{d}{d\lambda}\exp(\lambda\widehat{\xi})q\Big|_{\lambda=0} \;=\; \widehat{\xi}\,q \;=\; \xi \times q\,.$$

Thus, for each $\xi \in \mathbb{R}^3$, there is the conserved Noetherian momentum

$$J_\xi^L(q, \dot{q}) \;=\; \xi^Q \cdot p \;=\; (\xi \times q)m\dot{q} \;=\; \xi \cdot (mq \times \dot{q})\,.$$

Mechanically, this is (up to a factor) the component along the direction of ξ of the angular momentum of the system. By the arbitrariness of ξ, the vector $mq \times \dot{q}$ is conserved.

4. We consider now the two-body problem,

$$L(q_1, q_2, \dot{q}_1, \dot{q}_2) = \frac{1}{2}m_1|\dot{q}_1|^2 + \frac{1}{2}m_2|\dot{q}_2|^2 - \tilde{V}(|q_1 - q_2|)$$

in either $Q = \mathbb{R}^3 \times \mathbb{R}^3$ or $Q = \{(q_1, q_2) \in \mathbb{R}^3 \times \mathbb{R}^3 \,:\, q_1 \neq q_2\}$, and the 'diagonal' action Ψ of $\mathbb{R}^3$ of Q given by

$$\Psi_\xi(q_1, q_2) = (q_1 + \xi, q_2 + \xi)\,, \qquad \xi \in \mathbb{R}^3\,.$$

It is immediate to check that L is Ψ^{TQ} invariant. The infinitesimal generator associated to $\xi \in \mathbb{R}^3$ is the constant vector field

$$\xi^Q(q_1, q_2) = (\xi, \xi) \in \mathbb{R}^3 \times \mathbb{R}^3 = T_{(q_1,q_2)}Q\,.$$

Since the momentum covector is $p = \frac{\partial L}{\partial \dot{q}} = (m_1\dot{q}_1, m_2\dot{q}_2)$, the Noetherian momentum associated to $\xi \in \mathbb{R}^3$ is

$$\begin{aligned} J_\xi^L(q_1, q_2, \dot{q}_1, \dot{q}_2) \;&=\; p \cdot \xi^Q \\ &=\; (m_1\dot{q}_1, m_2\dot{q}_2) \cdot (\xi, \xi) \\ &=\; m_1\dot{q}_1 \cdot \xi + m_2\dot{q}_2 \cdot \xi \\ &=\; (m_1\dot{q}_1 + m_2\dot{q}_2) \cdot \xi \end{aligned}$$

namely, (up to a factor) the component along ξ of the total quantity of motion of the system $m_1\dot{q}_1 + m_2\dot{q}_2 \in \mathbb{R}^3$. Here too, because of the arbitrariness of ξ, we obtain the conservation of a vector.

Exercises 3.1.1 *(i) Show that in example 4. the Lagrangian is invariant under the lift of the diagonal action $\Psi_R(q_1, q_2) = (Rq_1, Rq_2)$ of SO(3) on Q and that this leads to the conservation of the momenta $(m_1q_1 \times \dot{q}_1 + m_2q_2 \times \dot{q}_2) \cdot \xi$, $\xi \in \mathbb{R}^3$, and hence of the angular momentum vector $m_1q_1 \times \dot{q}_1 + m_2q_2 \times \dot{q}_2$.*

3.1.C The momentum map. The fact that, in the previous examples, in presence of a symmetry group of dimension greater than one Noether theorem leads to the conservation of a *vector function* raises some issues. In particular, of which vector space is such conserved quantity a vector?

Recall that, by (2.1.9), for each $q \in Q$, the map $\xi \mapsto \xi^Q(q)$ is a linear map from $\mathfrak{g}$ to T_qQ. Therefore, for each $(q, \dot{q}) \in TQ$, the map

$$\xi \mapsto J^L_\xi(q, \dot{q}) \;=\; \big\langle p(q, \dot{q}), \xi^Q(q)\big\rangle \tag{3.1.6}$$

is a linear map from $\mathfrak{g}$ to $\mathbb{R}$. Since $\mathfrak{g} = T_eG$ is a vector space, for each $(q, \dot{q}) \in TQ$ the map (3.1.6) is an element of the dual space $\mathfrak{g}^*$. But then, there exists a map $J : TQ \to \mathfrak{g}^*$ which to each $(q, \dot{q})$ associates the map (3.1.6).

Definition 3.1.5 *Consider a Lagrangian system with Lagrangian $L : TQ \to \mathbb{R}$ and an action Ψ of a Lie group G on Q. The map*

$$J^L : TQ \to \mathfrak{g}^* , \qquad (q, \dot{q}) \mapsto J^L(q, \dot{q}) = \big(\xi \mapsto J^L_\xi(q, \dot{q})\big)$$

is the momentum map *of the lifted action Ψ^{TQ}.*

Note that

$$\big\langle J^L(q, \dot{q}), \xi\big\rangle \;=\; \big\langle p(q, \dot{q}), \xi^Q(q)\big\rangle \qquad \forall (q, \dot{q}) \in TQ\,,\ \xi \in \mathfrak{g} \tag{3.1.7}$$

where on the left of the equal sign the duality is that between $\mathfrak{g}^*$ and $\mathfrak{g}$ and on the right that between T^*Q and TQ.

Proposition 3.1.6 (Vector Noether theorem) *If $L : TQ \to \mathbb{R}$ is invariant under the tangent lift Ψ^{TQ} of an action of G on Q, then the momentum map of Ψ^{TQ} is constant along the solutions of the Euler-Lagrange equation for L.*

Proof. For each $\xi \in \mathfrak{g}$, $J^L(q, \dot{q})(\xi) = J^L_\xi(q, \dot{q})$ is constant along the solutions. ■

Thus, the 'Noetherian' first integrals of variational systems have in a natural way a vectorial character: they are vectors of the dual $\mathfrak{g}^*$ of the Lie algebra $\mathfrak{g}$ of the symmetry group. This is a vector space 'created by the group', not by the kinematic or dynamics of the system.

> **Example:** In all the examples of the previous section, the dual of the Lie algebra, as a vector space, is $\mathbb{R}^3$ and the momentum map takes values in $\mathbb{R}^3$.

If the momentum map $J^L : TQ \to \mathfrak{g}^*$ is conserved, then its level sets are invariant under the flow of Lagrange equations (seen as a first order system in TQ). Concerning their dimension:

Proposition 3.1.7 *If the Lagrangian L is regular and the action is free, then $J^L : TQ \to \mathfrak{g}^*$ is a submersion.*[1]

Proof. If L is regular (Remark (i) in section 3.1.A), then the map

$$\Lambda_L : TQ \to T^*Q\,, \qquad \Lambda_L(q,\dot q) := \big(q, p(q,\dot q)\big)$$

(the *Legendre transformation* induced by L) is a local diffeomorphism. Restrict to open sets $U \subseteq TQ$ and $U^* \subseteq T^*Q$ such that $\Lambda_L : U \to U^*$ is a diffeomorphism, and consider the map $J := J^L \circ \Lambda^{-1} : U^* \to \mathfrak{g}^*$. Since $\Lambda_L|_U$ is a diffeomorphism, $J^L : U \to \mathfrak{g}^*$ is a submersion if and only if $J : U^* \to \mathfrak{g}^*$ is a submersion. From $J^L(q,\dot q) = J\big(q, p(q,\dot q)\big)$ and (3.1.7) it follows that

$$\big\langle J(q,p), \xi \big\rangle = \big\langle p, \xi^Q(q) \big\rangle \qquad \forall (q,p) \in U^*\,,\ \xi \in \mathfrak{g}\,.$$

Choose a basis $\xi_1, \dots, \xi_h$, $h = \dim G$, of $\mathfrak{g}$ and consider the dual basis $\xi_1^*, \dots, \xi_h^*$ of $\mathfrak{g}^*$. Thus

$$J(q,p) \;=\; \sum_{i=1}^h \langle J(q,p), \xi_i \rangle \xi_i^* \;=\; \sum_{i=1}^h \big\langle p, \xi_i^Q(q) \big\rangle \xi_i^*\,. \tag{3.1.8}$$

Saying that $(q,p) \mapsto J(q,p)$ is a submersion at (q,p) means that it has rank h. It is thus sufficient to prove that, for each q, the linear map $p \mapsto J(q,p)$ from T_q^*Q to $\mathfrak{g}^*$ has rank h. The last expression in (3.1.8) shows that this happens if the linear map

$$p \mapsto \big(\big\langle p, \xi_1^Q(q) \big\rangle, \dots, \big\langle p, \xi_h^Q(q) \big\rangle\big)$$

from T_qQ^* to $\mathbb{R}^h$ has rank h, and in turn this happens if the vectors $\xi_1^Q(q), \dots, \xi_h^Q(q)$ are linearly independent. This follows from the hypothesis that the action Ψ on Q is free because $\xi_i^Q(q) = T_e\Psi^q \cdot \xi_i$ (see (2.1.9)) and, if Ψ is free, then the proof of Proposition 2.1.10 shows that $T_e\Psi^q$ is injective. ∎

Remarks: (i) Noether theorem, and the momentum map, extend to actions of Lie groups that act (in a suitable) way on the tangent bundle TQ and are not the lifts of actions on Q. An important example is the SO(4) symmetry of the spatial Kepler system, whose momentum map produces the angular momentum and the Laplace-Runge-Lenz vector. A thorough understanding of the momentum map of non-lifted actions requires however the theory of Hamiltonian systems on symplectic manifolds.

(ii) The momentum map of Definition (3.1.5) depends on the group action and on the Lagrangian. If the Lagrangian is such that the Legendre transformation Λ_L is a diffeomorphism, then the system can be studied in the Hamiltonian formulation. The Legendre transformation conjugates the Lagrange equations

[1] In class I did not prove this fact and I forgot mentioning the hypothesis that L must be regular.

for L to the Hamilton equations for the Hamiltonian $H = (\sum_i p_i \dot q_i - L) \circ \Lambda_L^{-1}$. In the Hamiltonian formulation, the momenta p_i are coordinates, and the momentum map depends only on the group action. There is a theory of Lie group actions on symplectic manifolds that extends the Hamiltonian analogue of lifted actions considered here and leads to momentum maps which depend only on the group actions and are conserved for all Hamiltonian systems with invariant Hamiltonian.

Exercises 3.1.2 *(i) Show that if an action Ψ is free then its tangent lift is free.*

3.1.D Routh reduction. The invariance of the Lagrangian under a lifted action not only produces the conservation of the momentum map but, as we now see, may also be used to reduce the Lagrange equations, and the two operations may to a certain extent be combined.

We are interested to the case of regular Lagrangians $L : TQ \to \mathbb{R}$, so that Lagrange equations for L can be viewed as a vector field X^L on TQ.

Proposition 3.1.8 *Assume that a regular Lagrangian $L : TQ \to \mathbb{R}$ is invariant under the lift Ψ^{TQ} of an action Ψ on Q. Then, X^L (namely, the Lagrange equations for L) is invariant under $T\Psi^{TQ}$.*

Proof. It is known from the introductory courses on Lagrangian mechanics that if $\mathcal{C} : Q \to Q$ is a diffeomorphism, then $T\mathcal{C} : TQ \to TQ$ conjugates X^L to $X^{L\circ T\mathcal{C}}$ ('geometric invariance' or 'invariance in form' of Lagrange equations). If L is Ψ^{TQ}-invariant, then $(T\Psi_g)_* X^L = X^{L\circ T\Psi_g} = X^L$ for all $g \in G$. ■

We note that the invariance of the Lagrangian is stronger than the invariance of Lagrange equations:

Example: Equation
$$\ddot x = -g\,, \qquad x \in \mathbb{R}\,,$$
namely the first order system $\dot x = v$, $\dot v = -g$ in $T\mathbb{R} \ni (x, v)$, is invariant under the lifted action $\Psi_\lambda^{TQ}(x, v) = (x + \lambda, v)$ of $\mathbb{R} \ni \lambda$. This equation is the Lagrange equation of the Lagrangian
$$L(x, \dot x) = \frac{1}{2}\,\dot x^2 - gx$$
which however is *not* invariant under Ψ^{TQ} because $L(x + \lambda, \dot x) = L(x, \dot x) - g\lambda$. In this case there is a symmetry of the Lagrange equation, which can be used to reduce the system (see Exercise 2.2.4.i), but which does not produce a first integral.

The possibility of performing both operations—first restrict to a level set of the momentum map and then reduce under the group action—is at the basis of the classical integration methods in Lagrangian mechanics. However, in general the level sets of the momentum map are not invariant under the action of the entire group, but only of certain subgroups of it, and reduction must be restricted to the action of these subgroups.

Recall that J is a map from TQ to the dual $\mathfrak{g}^*$ of the Lie algebra of G, and that G acts on $\mathfrak{g}^*$ with the coadjoint action (section 2.1.B, example 12).

Proposition 3.1.9 *Let Ψ be an action of a Lie group G on Q and $L : TQ \to \mathbb{R}$ be a Ψ^{TQ}-invariant Lagrangian. Then*

$$J^L \circ T\Psi_g = (\mathrm{Ad}^*)_g \circ J^L \qquad \forall g \in G \tag{3.1.9}$$

(namely, J^L is equivariant with respect to the action Ψ^{TQ} on TQ and to the coadjoint action Ad^* on $\mathfrak{g}^*$).

Proof. (3.1.9) is equivalent to $\big\langle J^L\big(T\Psi_g(q,\dot q)\big),\xi\big\rangle = \big\langle(\mathrm{Ad}^*)_g\big(J^L(q,\dot q)\big),\xi\big\rangle$ for all $\xi\in\mathfrak{g}$ and $(q,\dot q)\in TQ$. Since $(\mathrm{Ad}^*)_g$ is the adjoint map of $\mathrm{Ad}_{g^{-1}}$, see (2.1.4), (3.1.9) is also equivalent to $\big\langle J^L\big(T\Psi_g(q,\dot q)\big),\xi\big\rangle = \big\langle J^L(q,\dot q),\mathrm{Ad}_{g^{-1}}\xi\big\rangle$ for all ξ and $(q,\dot q)$. In turn, this is equivalent to $J^L_\xi\big(T\Psi_g(q,\dot q)\big) = J^L_{\mathrm{Ad}_{g^{-1}}\xi}(q,\dot q)$ for all ξ and $(q,\dot q)$, namely to

$$J^L_\xi \circ T\Psi_g = J^L_{\mathrm{Ad}_{g^{-1}}\xi} \qquad \forall g\in G\,,\ \xi\in\mathfrak{g}\,, \tag{3.1.10}$$

which is what we prove.

Fix $g\in G$ and denote $p^L = \frac{\partial L}{\partial\dot q}$ and $p^{L\circ T\Psi_g} = \frac{\partial}{\partial\dot q}(L\circ T\Psi_g)$ the momenta covectors of the two Lagrangians L and $L\circ T\Psi_g$, respectively. Thus, $J^L_\xi = \big\langle p^L,\xi^Q\big\rangle$ and $J^{L\circ T\Psi_g}_\xi = \big\langle p^{L\circ T\Psi_g},\xi^Q\big\rangle$ for all $\xi\in\mathfrak{g}$. For each $i=1,\dots,d$,

$$\begin{aligned}
p_i^{L\circ T\Psi_g}(q,\dot q) &= \frac{\partial}{\partial\dot q_i}L\big(\Psi_g(q),\Psi_g'(q)\dot q\big)\\
&= \sum_j \frac{\partial L}{\partial\dot q_j}\big(\Psi_g(q),\Psi_g'(q)\dot q\big)\frac{\partial}{\partial\dot q_i}\big(\Psi_g'(q)\dot q\big)_j\\
&= \sum_j\Big(\frac{\partial L}{\partial\dot q_j}\circ T\Psi_g(q,\dot q)\Big)\Psi_g'(q)_{ji}\\
&= \sum_j\big(p_j^L\circ T\Psi_g(q,\dot q)\big)\big(\Psi_g'(q)^T\big)_{ij}
\end{aligned}$$

and thus

$$p^{L\circ T\Psi_g} = (p^L\circ T\Psi_g)\Psi_g'^T\,.$$

Therefore

$$\begin{aligned}
J_\xi^{L\circ T\Psi_g} &= \big\langle p^{L\circ T\Psi_g},\xi^Q\big\rangle\\
&= \big\langle (p^L\circ T\Psi_g)\Psi_g'^T,\xi^Q\big\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \left\langle p^L \circ T\Psi_g , \Psi'_g \xi^Q \right\rangle \\
&= \left\langle p^L \circ T\Psi_g , (\Psi_g)_* \xi^Q \circ \Psi_g \right\rangle \\
&= \left\langle p^L , (\Psi_g)_* \xi^Q \right\rangle \circ T\Psi_g
\end{aligned}
$$

Recalling $(\Psi_g)_*\xi^Q = (\mathrm{Ad}_g\xi)^Q$ from proposition 2.1.9, this gives

$$
J_\xi^{L\circ T\Psi_g} \;=\; \left\langle p^L, (\mathrm{Ad}_g\xi)^Q \right\rangle \circ T\Psi_g \;=\; J^L_{\mathrm{Ad}_g\xi} \circ T\Psi_g\,.
$$

Thus, if L is Ψ^{TQ}-invariant, $J^L_\xi = J^L_{\mathrm{Ad}_g\xi} \circ T\Psi_g$ for all ξ. Formula (3.1.10) is obtained changing g into g^{-1} and using $(T\Psi_{g^{-1}})^{-1} = T(\Psi^{-1}_{g^{-1}}) = T\Psi_g$. ■

Example: Consider the SO(3)-invariant Lagrangian $L(q,\dot q) = \frac{1}{2}m|\dot q|^2 - V(|q|)$, $q \in \mathbb{R}^3$ (example 3, section 3.1.B). The action is the linear action of SO(3) on $\mathbb{R}^3$. If we identify the Lie algebra of SO(3) with $\mathbb{R}^3$ then $J^L_\xi(q,\dot q) = \xi \cdot (mq \times \dot q)$ for all $\xi \in \mathbb{R}^3 \equiv \mathfrak{so}(3)$. In turn, if we identify $\mathfrak{so}(3)^*$ with $\mathbb{R}^3$ by means of the standard inner product of $\mathbb{R}^3 \equiv \mathfrak{so}(3)$, this gives

$$
J^L(q,\dot q) = mq \times \dot q\,.
$$

The lifted action is $R.(q,\dot q) = (Rq, R\dot q)$. Note that, if $R \in \mathrm{SO}(3)$, then $(Ru) \times (Rv) = R(u \times v)$ for all $u, v \in \mathbb{R}^3$. Thus, for any $R \in \mathrm{SO}(3)$,

$$
J^L(Rq, R\dot q) = m(Rq) \times (R\dot q) = mR(q \times \dot q) = RJ^L(q,\dot q)
$$

and $(Ad^*)_R = R$ (see example 13, section 3.1.B).

Corollary 3.1.10 *In the hypotheses of Proposition 3.1.9, for each $\mu \in J^L(TQ) \subseteq \mathfrak{g}^*$, the level set $(J^L)^{-1}(\mu)$ is invariant under the restriction of the lifted action Ψ^{TQ} to $G^*_\mu \times TQ$, where G^*_μ is the isotropy subgroup of μ in the coadjoint action of G on $\mathfrak{g}^*$.*

Proof. If $(q,\dot q) \in (J^L)^{-1}(\mu)$ then, by (3.1.9), for all $g \in G^*_\mu$, $J^L\big(T\Psi_g(q,\dot q)\big) = (\mathrm{Ad}_g)^*(J^L(q,\dot q)) = (\mathrm{Ad}_g)^*(\mu) = \mu$. ■

Under the appropriate hypotheses, therefore, if the Lagrangian is invariant under a free lifted action, the system can first be restricted to each level set $J^L = \mu$ of the momentum map, which has dimension $2\dim Q - \dim G$, and then reduced under the action of the isotropy subgroup G^*_μ. This produces a μ-parametrized family of reduced systems on phase spaces of dimension $2\dim Q - \dim G - \dim G^*_\mu$. Much more can be said about them. We just remark that this is a generalization of the Routh reduction (or ignoration of coordinates) studied in the courses on Lagrangian mechanics, which is that in which G is abelian. The coadjoint action for an abelian group is trivial, namely $(Ad^*)_g = \mathrm{id}$ for all g (simple exercise), and therefore $G^*_\mu = G$. In this case, all reduced spaces have dimension $2\dim Q - 2\dim G$. A deeper analysis shows that the reduced phase space is the tangent bundle of Q/G and each reduced system is still a Lagrangian system.

Chapter A

Appendix: Vector fields

The aim of this appendix is to review a few basic facts about vector fields. We assume some basic facts as known, in particular the notion of (smooth) manifold.

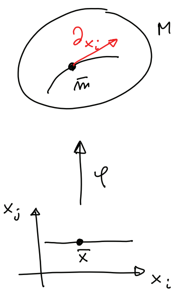
We will occasionally use the following terminology. Let M be an n-dimensional manifold. Let $\psi : U_M \to U$, with open sets $U_M \subseteq M$ and $U \subseteq \mathbb{R}^n$ is a local chart, or local coordinate map, of M. Then ψ is a diffeomorphism. Its components $(x_1, \ldots, x_n)$ are said to be the coordinate functions, or coordinates. The inverse $\varphi := \psi^{-1} : U \to U_M$ of the chart map ψ is a diffeomorphism and will be called local parametrization. Since we will use φ more than ψ, we will denote the latter φ^{-1}.

A.1 Vector fields and differential forms

A.1.A The tangent and cotangent bundles. In the following, M is a (smooth) manifold of dimension n.

The *tangent vectors* to M at a point $m \in M$ are the derivatives $\frac{d}{dt}\gamma(t)|_{t=0}$ of (smooth) curves $\gamma : \mathbb{R} \to M$ that pass through m, namely $\gamma(0) = m$.

At each point $m \in M$, tangent vectors to M at m form a vector space $T_m M$ of dimension $n = \dim M$, the *tangent space* to M at m. As a set, the *tangent bundle* TM of M is the disjoint union of all tangent spaces to M. The smooth structure of M induces a smooth structure on TM, which makes it a smooth manifold of dimension $2n$. This structure is given by an atlas whose charts are build as follows.



Consider a local parametrization $\varphi : U \to U_M$ of M, with coordinates $(x_1, \ldots, x_n) = \varphi^{-1}$. For each $m = \varphi(x) \in U_M$, with $x \in U$, the n *coordinate curves* through m are the curves $\lambda \mapsto \varphi(x + \lambda e_i)$, $i = 1, \ldots, n$, with e_i the i-th vector of the canonical basis of $\mathbb{R}^n$. The derivatives

$$\partial_{x_i}(m) \ := \ \frac{d}{d\lambda}\varphi(x+\lambda e_i)\Big|_{\lambda=0}$$

of these curves are tangent vectors to M at m and, since φ is a diffeomorphism, are linearly independent. Therefore, they form a basis for T_mM, called the *natural basis* induced by the considered parametrization (or by the coordinate system $\varphi^{-1} : U_M \to U$). We may therefore define a local parametrization $T\varphi : U \times \mathbb{R}^n \to TU_M$ for TU_M as follows: the point of TU_M of coordinates $(x, v) \in U \times \mathbb{R}^n$ is[1]

$$(T\varphi)(x,v) := \sum_i v_i \partial_{x_i}(\varphi(x)) \in T_{\varphi(x)}M\,.$$

In this way, we build an atlas for TM, which is called a "bundle atlas" (and its local coordinates are called 'bundle' or 'lifted' coordinates).

For each $m \in M$, the *cotangent space* T^*M to M is the dual of T_mM (namely, the space of linear forms on T_mM). The *cotangent bundle* T^*M of M is the disjoint union of all its cotangent spaces T^*_mM. Like TM, also T^*M has a smooth structure induced by that of M. Consider a local parametrization $\varphi : U \to U_M$ of M. In each cotangent space T^*_mM, $m \in U_M$, consider the 'dual basis' to the natural basis $\partial_1(m), \ldots, \partial_n(m)$ of T_mM. This basis is formed by the covectors

$$\partial^*_{x_1}(m), \ldots, \partial^*_{x_n}(m) \in T^*_mM$$

defined by

$$\langle \partial^*_{x_i}(m), \partial_{x_j}(m)\rangle_m = \delta_{ij} \qquad \forall i, j$$

where $\langle \ , \ \rangle_m$ denotes the pairing between T^*_mM and T_mM. (Later on, for reasons which will be explained, the covectors $\partial^*_i(m)$ will be denoted $dx_i(m)$). In this way we obtain on T^*M local 'bundle' parametrizations $T^*\varphi : U \times \mathbb{R}^n \to T^*U_M$ defined as follows: the point of T^*M of coordinates (x, α) is $\sum_i \alpha_i \partial^*_i(\varphi(x)) \in T^*_{\varphi(x)}M$.

Exercises A.1.1 *(i) Show that the transition function between two local bundle parametrizations $T\varphi$ and $T\tilde\varphi$ of TM is $(x, v) \mapsto (\tilde x, \tilde v) = (\mathcal{C}(x), \mathcal{C}'(x)v)$ with $\mathcal{C} = \tilde\varphi^{-1} \circ \varphi$.*

(ii) Verify that, if $\Theta = \sum \alpha_i \partial^*_i \in T^*_mM$ and $V = \sum v_i\partial_i \in T_mM$ then $\langle \Theta, V\rangle = \sum_i \alpha_i v_i$.

(iii) Show that the transition function between two local bundle parametrizations $T^*\varphi$ and $T^*\tilde\varphi$ of TM is $(x, \alpha) \mapsto (\tilde x, \tilde\alpha) = \big(\mathcal{C}(x), \mathcal{C}'(x)^{-T}\alpha\big)$ with $\mathcal{C} = \tilde\varphi^{-1} \circ \varphi$.

[1] At this stage, $T\varphi$ should be considered as a symbol defined by this condition. Later on it will be clear that it is in fact the tangent map of φ.

A.1.B Vector fields and differential forms. A *vector field* on a manifold M is a (smooth) map $X : M \to TM$ which to each point $m \in M$ associates a vector $X(m) \in T_mM$. Similarly, a *differential* 1*-form* or simply a *1-form* on M is a (smooth) map $\Theta : M \to T^*M$ with the property that $\Theta(m) \in T^*_mM$ for all $m \in M$. We will denote by $\mathcal{X}(M)$ the set of all vector fields on M and by $\mathcal{X}^*(M)$ the set of all 1-form on M. They are both (infinite-dimensional) vector spaces.

Globalizing the pairing $\langle\ ,\ \rangle_m$ between tangent and cotangent spaces we obtain a pairing $\langle\ ,\ \rangle$ between vector fields and 1-forms, namely a bilinear map

$$\langle\ ,\ \rangle : \mathcal{X}^*(M) \times \mathcal{X}(M) \ \to\ C^\infty(M) \tag{A.1.1}$$

which associates to each 1-form Θ and vector field X a real function $\langle \Theta, X\rangle$ on M defined as

$$\langle \Theta, X\rangle(m) := \langle \Theta(m), X(m)\rangle_m\,, \quad m \in M\,.$$

Local representatives. Consider a local parametrization $\varphi : U \to U_M$ of M with coordinates $(x_1, \ldots, x_n) = \varphi^{-1}$. The tangent vectors to the coordinate lines define vector fields $\partial_1, \ldots, \partial_n$ in $U_M = \varphi(U)$, whose values at each point $m \in U_m$ form the natural basis of T_mM. We thus define the local representative $X^{\rm loc} = (X^{\rm loc}_1, \ldots, X^{\rm loc}_n) : U \to \mathbb{R}^n$ of a vector field X on M as the vector field via

$$X|_{U_M} = \sum_{i=1}^n (X^{\rm loc}_i \circ \varphi^{-1})\partial_{x_i}\,.$$

Similarly, the local representative $\Theta^{\rm loc} = (\Theta^{\rm loc}_1, \ldots, \Theta^{\rm loc}_n) : U \to \mathbb{R}^n$ of a 1-form Θ is defined by

$$\Theta|_{U_M} = \sum_{i=1}^n (\Theta^{\rm loc}_i \circ \varphi^{-1})\partial^*_{x_i}\,.$$

Exercises A.1.2 *(i) Let $\varphi^{-1} : U_M \to U$ and $\tilde\varphi^{-1} : \tilde U_M \to U$ be two local coordinate systems in M with non-disjoint domains. Show that the local representatives $X^{\rm loc}$, $\tilde X^{\rm loc}$ of a vector field X and the local representatives $\Theta^{\rm loc}$, $\tilde\Theta^{\rm loc}$ of a 1-form Θ are related by*

$$\tilde X^{\rm loc}(\mathcal{C}(x)) = \mathcal{C}'(x)X(x)\,, \qquad \tilde\Theta(\mathcal{C}(x)) = \mathcal{C}'(x)^{-T}\Theta(x) \qquad \forall\, x \in U$$

with (now: this is the inverse of the $\mathcal{C}$ of Exercise A.1.1.iii) $\mathcal{C} = \varphi^{-1} \circ \tilde\varphi$.

(ii) Verify that if $X|_{U_M} = \sum_{i=1}^n (X^{\rm loc}_i \circ \varphi^{-1})\partial_{x_i}$ and $\Theta|_{U_M} = \sum_{i=1}^n (\Theta^{\rm loc}_i \circ \varphi^{-1})dx_i$, then $\langle\Theta, X\rangle_m \circ \varphi = \sum_i X^{\rm loc}_i \Theta^{\rm loc}_i$ for all $m \in U_M$.

A.1.C Tangent maps and differentials (exterior derivatives). Every (smooth) map $\Psi : M \to N$ between two manifolds induces a linear map

$$T_m\Psi : T_mM \to T_{\Psi(m)}N$$

from each tangent space T_mM to M and the tangent space to N at the point $\Psi(m)$. This map, which plays the role of derivative, is called the *tangent map of* Ψ *at* m and is defined as follows. Recall that tangent vectors are derivatives of curves. If $v = \gamma'(0) \in T_mM$ with a curve $\gamma : \mathbb{R} \to M$, $\gamma(0) = m$, then $T_m\Psi(v)$ is defined as the derivative at $t = 0$ of the transformed curve $\Psi \circ \gamma : \mathbb{R} \to N$, namely:

$$T_m\Psi\big(\gamma'(0)\big) \;\mapsto\; (\Psi \circ \gamma)'(0)\,.$$

Note that $(\Psi \circ \gamma)(0) = \Psi(m)$ and hence $(\Psi \circ \gamma)'(0) \in T_{\Psi(m)}N$. Globalizing this construction, we obtain a map

$$T\Psi : TM \to TN$$

which is defined by $T\Psi|_{T_mM} := T_m\Psi$ for any $m \in M$. The map $T\Psi$ can be shown to be smooth and is called the *tangent map of* Ψ.

Let us now specialize to the case $N = \mathbb{R}$. At each point $m \in M$, the tangent map $T_m\mathcal{F}$ of a function $\mathcal{F} : M \to \mathbb{R}$ is a linear map from T_mM to $T_{\mathcal{F}(m)}\mathbb{R} = \mathbb{R}$ and, as such, can be regarded as a covector $d\mathcal{F}(m) \in T_m^*M$, which is defined by

$$T_m\mathcal{F} \cdot v = \langle d\mathcal{F}(m), v\rangle_m \qquad \forall\, v \in T_mM\,. \tag{A.1.2}$$

Thus, a function $\mathcal{F} : M \to \mathbb{R}$ defines a differential 1-form

$$d\mathcal{F} \in \mathcal{X}^*(M)\,,$$

called the *differential* or the *exterior derivative* of $\mathcal{F}$. The map $d : C^\infty(M) \to \mathcal{X}^*(M)$ so defined is called *differential* or *exterior derivative* of functions.

Note that, in particular, in a chart $\varphi^{-1} = (x_1, \dots, x_n) : U_M \to U \subseteq \mathbb{R}^n$ of M, the differential of the i-th coordinate function $x_i : U_M \to \mathbb{R}$ at a point $m \in U_M$ is

$$dx_i(m) = \partial^*_{x_i}(m)\,.$$

(Indeed, at a point $m = \varphi(x)$, $\langle dx_i(m), \partial_{x_j}(m)\rangle_m = T_mx_i \cdot \partial_{x_j}(m) = T_mx_i \cdot \frac{d}{dt}\varphi(x+te_j)|_{t=0} = \frac{d}{dt}x_i \circ \varphi(x+te_j)|_{t=0} = \delta_{ij}$ because $x_i \circ \varphi(x+te_j)$ is the i-th component of $x + te_j$ and equals $x_i + t$ if ij and x_i if $i \neq j$).

Because of this, in each chart, the vectors $\partial^*_{x_1}(m), \dots, \partial^*_{x_n}(m)$ of the natural basis of T_m^*M are denoted $dx_1(m), \dots, dx_n(m)$. Hence,

$$\Theta(m) = \sum_i \Theta_i^{\mathrm{loc}}(x) dx_i(\varphi(x)) \qquad \forall x \in U\,.$$

Exercises A.1.3 *(i) Show that if Ψ_{loc} is the local representative of a map $\Psi : M \to N$ relative to two charts of M and N, then the local representative of $T_m\Psi$ in the associated bundle charts of TM and TN is $(x, v) \mapsto (\Psi_{\mathrm{loc}}(x), \Psi'_{\mathrm{loc}}(x)v)$, where $\Psi'_{\mathrm{loc}}(x)$ is the Jacobian matrix of Ψ_{loc} at the point x.*

(ii) Show that, if $\mathcal{F}_{\mathrm{loc}}$ is the local representative of $\mathcal{F} : M \to \mathbb{R}$ in a chart $\varphi^{-1} = (x_1, \dots, x_n) : U_M \to U \subseteq \mathbb{R}^n$ of M, then the local representative of $d\mathcal{F}(\varphi(x))$, thought of as an element

of $T^*_m M$, is $\mathcal{F}'_{\rm loc}(x) = \left(\frac{\partial \mathcal{F}_{\rm loc}}{\partial x_1}(x), \ldots, \frac{\partial \mathcal{F}_{\rm loc}}{\partial x_n}(x)\right)$, which in calculus is called the differential $d\mathcal{F}_{\rm loc}(x)$ of $\mathcal{F}_{\rm loc}$ at x.

(iii) Show that, if Ψ is a diffeomorphism, then so is $T\Psi$ and, in coordinates (see (i) above) it is the map $(y,w) \mapsto (\Psi^{-1}_{\rm loc}(y), (\Psi^{-1}_{\rm loc})'(y)w) = (\Psi^{-1}_{\rm loc}(y), \Psi'_{\rm loc}(\Psi^{-1}_{\rm loc}(y))w)$. [Hint: for the last expression use the chain rule.]

A.1.D Flows. Vector fields have two roles—or a double nature: they can be integrated to define flows, and they act as derivations (of functions, vector fields, differential forms and in fact all tensor objects). We begin with the first aspect.

An *integral curve* of X is a smooth curve

$$\gamma : I \to M\,, \quad t \mapsto \gamma(t)$$

where $I \subset \mathbb{R}$ is an interval, such that

$$\frac{d\gamma}{dt}(t) \;=\; X(\gamma(t)) \qquad \forall t \in I\,.$$

If $\gamma(0) = m$, then we will say that γ is the integral curve "through m", or which "passes through m" (understood: at $t = 0$). The existence, and (up to the choice of I) the uniqueness, of an integral curve through any point is granted by the existence and uniqueness theorem for ODEs.

We will assume (usually tacitly) that all vector fields we consider are *complete*, namely, that their integral curves can be defined for all times, in the precise sense that they all have (maximal) existence interval $I = \mathbb{R}$.

If X is complete, then there exists its *flow*

$$\Phi^X : \mathbb{R} \times M \to M$$

which is defined by the fact that $\Phi^X(t,m)$ is the value at time t of the integral curve of X which passes through m at $t = 0$. For each $t \in \mathbb{R}$, the "map at time t of the flow" is the map

$$\Phi^X_t : M \to M\,, \quad \Phi^X_t(m) = \Phi^X(t,m)\,.$$

The definition of Φ^X is thus equivalent to the two following conditions:

$$\Phi^X_0 = {\rm id}_M\,, \qquad \tfrac{d}{dt}\Phi^X_t = X \circ \Phi^X_t \qquad \forall t \in \mathbb{R}\,. \tag{A.1.3}$$

The smooth dependence of solutions of ODEs from the initial datum implies that, for each t, Φ^X_t is a smooth map. Moreover, the 'time translatability' of the solutions of (autonomous) ODEs implies that

$$\Phi^X_{t+s} \;=\; \Phi^X_t \circ \Phi^X_s \qquad \forall t,s \in \mathbb{R}\,. \tag{A.1.4}$$

This implies that, for any t, Φ^X_t is invertible and its inverse is Φ^X_{-t}. Thus, for each t, Φ^X_t is a diffeomorphism of M onto itself.

Remarks: (i) All vector fields on a compact manifolds are complete.

(ii) (A.1.4) means that the flow of a vector field can be regarded as a group homomorphism from the abelian group $\mathbb{R}$ into the infinite-dimensional group of all diffeomorphisms of M, with the composition as product.

A.1.E Push-forward and pull-back of functions and vector fields.

Diffeomorphisms (and sometimes smooth maps) $\Psi : M \to N$ can be used to transport geometric objects from one manifold to the other.

$$\begin{tikzcd} M \arrow[rr, "\Psi"] \arrow[dr, "f"'] & & N \arrow[dl, squiggly, "\Psi_* f"] \\ & \mathbb{R} & \end{tikzcd} .$$

$$\begin{tikzcd} M \arrow[rr, "\Psi"] \arrow[dr, squiggly, "\Psi^* g"'] & & {} \arrow[dl, "g"] \\ & \mathbb{R} & \end{tikzcd}$$

Definition A.1.1 *The* push-forward *of functions under a diffeomorphism $\Psi : M \to N$ is the map*

$$\Psi_* \ : \ C^\infty(M) \to C^\infty(N) \, , \qquad \Psi_* f := f \circ \Psi^{-1}$$

The pull-back *of functions under a (smooth) map $\Psi : M \to N$ is the map*

$$\Psi^* \ : \ C^\infty(N) \to C^\infty(M) \, , \qquad \Psi^* g := g \circ \Psi \, .$$

The reason why the pull-back of functions can be defined for maps that need not be diffeomorphisms is that it does not involve the inverse of the map.

$$\begin{tikzcd} M \arrow[r, "\Psi"] \arrow[d, "X"'] & N \arrow[d, squiggly, "\Psi_* X"] \\ TM \arrow[r, "T\Psi"] & TN \end{tikzcd}$$

$$\begin{tikzcd} M \arrow[r, "\Psi"] \arrow[d, squiggly, "\Psi^* Y"'] & N \arrow[d, "Y"] \\ TM \arrow[r, "T\Psi"] & TN \end{tikzcd}$$

Definition A.1.2 *Let $\Psi : M \to N$ be a diffeomorphism. The* push-forward $\Psi_* \ : \ \mathfrak{X}(M) \to \mathfrak{X}(N)$ *and the* pull-back $\Psi^* \ : \ \mathfrak{X}(M) \to \mathfrak{X}(N)$ *of vector fields under a diffeomorphism $\Psi : M \to N$ are defined as*

$$\Psi_* X := (T\Psi \cdot X) \circ \Psi^{-1} \, , \qquad \Psi^* Y := ((T\Psi)^{-1} \cdot X) \circ \Psi \, .$$

Here, the inverse of either Ψ or $T\Psi$ appear in both cases; thus, push-forward and pull-back of vector fields are defined only for diffeomorphisms.

In detail, if $X \in \mathfrak{X}(M)$, then

$$(\Psi_* X)(n) = T_{\Psi^{-1}(n)} \Psi \cdot X\big(\Psi^{-1}(n)\big) \qquad \forall n \in N \, ,$$

and similarly $(\Psi^* Y)(m) = (T_m \Psi)^{-1} \cdot Y(\Psi(m))$ for all $m \in M$. Note also that

$$(\Psi_* X)(\Psi(m)) \ = \ T_m \Psi \cdot X(m) \qquad \forall \, m \in M$$

Push-forward and pull backs of vector fields have a dynamical meaning:

$$\begin{tikzcd} M \arrow[r, "\Psi"] \arrow[d, "\Phi_t^X"'] & M \arrow[d, "\Phi_t^{\Psi_* X}"] \\ M \arrow[r, "\Psi"] & M \end{tikzcd}$$

Proposition A.1.3 *Let $\Psi : M \to M$ be a diffeomorphism and $X \in \mathfrak{X}(M)$. Then $\Psi_* X$ is the (unique) vector field on M whose integral curves are the images under Ψ of those of X. Equivalently, if $Z \in \mathfrak{X}(M)$, then*

$$\Psi \circ \Phi_t^X = \Phi_t^Z \circ \Psi \quad \forall t \in \mathbb{R} \quad \iff \quad Z = \Psi_* X \, .$$

Proof. If $t \mapsto \gamma(t) \in M$ is a curve, then

$$\gamma' = X \circ \gamma \quad \iff \quad (\Psi \circ \gamma)' = (\Psi_* X) \circ (\Psi \circ \gamma) \, .$$

Thus, Ψ maps integral curves of X into integral curves of $\Psi_* X$ and this uniquely defines $\Psi_* X$ (a vector field is uniquely determined by its integral curves). The equivalence between the two statements follows from the fact that $t \mapsto \Phi_t^X(m)$ is the integral curve of X through m etc. ■

A.1.F Push-forward and pull-back of differential 1-forms. Push-forward and pull-back of differential 1-forms are defined so as to be consistent, via the pairing (A.1.1), with those of functions and vector fields. Specifically, the *push-forward* under a diffeomorphism $\Psi : M \to N$ of a 1-form Θ on M is the 1-form $\Psi_*\Theta$ on N defined by

$$\langle \Psi_*\Theta, Y\rangle := \Psi_*\langle \Theta, \Psi^*Y\rangle = \langle \Theta, \Psi^*Y\rangle \circ \Psi^{-1} \qquad \forall\, Y \in \mathfrak{X}(N) \qquad \text{(A.1.5)}$$

$$\begin{tikzcd} M \arrow[dr, "\langle\Theta{,}\Psi^*Y\rangle"'] & & N \arrow[ll, "\Psi^{-1}"'] \arrow[dl, squiggly, "\langle\Psi_*\Theta{,}Y\rangle"] \\ & \mathbb{R} & \end{tikzcd}$$

where the second equality follows from the definition of push-forward of functions.[2] Similarly, the *pull-back* under Ψ of a 1-form Ξ on N is the 1-form $\Psi^*\Xi$ on M defined by

$$\langle \Psi^*\Xi, X\rangle := \Psi^*\langle \Xi, \Psi_*X\rangle = \langle \Xi \circ \Psi, T\Psi \cdot X\rangle \qquad \forall\, X \in \mathfrak{X}(M)\,. \qquad \text{(A.1.6)}$$

$$\begin{tikzcd} M \arrow[rr, "\Psi"] \arrow[dr, squiggly, "\langle\Psi^*\Xi{,}X\rangle"'] & & N \arrow[dl, "\langle\Xi{,}\Psi_*X\rangle"] \\ & \mathbb{R} & \end{tikzcd}$$

The last expression (which follows from the definitions of pull-back of functions and of push-forward of vector fields) shows that the pull-back of 1-forms, as that of functions, is defined for any (smooth) map, not only for diffeomorphisms.

Proposition A.1.4 *If $\Psi : M \to N$ is a (smooth) map then*

$$\Psi^*(dg) \;=\; d(\Psi^*g) \qquad \forall g \in C^\infty(N)$$

and if it is a diffeomorphism then

$$\Psi_*(df) \;=\; d(\Psi_*f) \qquad \forall f \in C^\infty(M)\,.$$

Proof. Take $m \in M$ and $v \in T_mM$. Then, using (A.1.2), $\langle d(\Psi^*f)(m), v\rangle_m = T_m(\Psi^*f) \cdot v = T_m(f \circ \Psi) \cdot v = T_{\Psi(m)}f \cdot T_m\Psi \cdot v$ and using (A.1.6), $\langle (\Psi^*df)(m), v\rangle_m = \langle df(\Psi(m)), T_m\Psi \cdot v\rangle_m = T_{\Psi(m)}f \cdot T_m\Psi \cdot v$. The formula for the push-forward is proven analogously. ■

The identities of Proposition A.1.4 are referred to as the 'naturalness' of the exterior differential (of functions) with respect to pull-back and push-forward.

Exercises A.1.4 *(i) Show that push-forward and pull-back, both of functions and of vector fields, are linear isomorphisms and are the inverse of each other.*

(ii) Show that for any map $\Psi : M \to N$ and 1-form $\Xi \in \mathfrak{X}^*(N)$,

$$(\Psi^*\Xi)(m) \;=\; (T_m\Psi)^*\,\Xi(\Psi(m))$$

where $(T_m\Psi)^* : T^*_{\Psi(m)}N \to T^*_mM$ is the adjoint of the linear map $T_m\Psi : T_mM \to T_{\Psi(m)}N$.

(iii) Verify that, in bundle coordinates on T^*M, the local representative of the adjoint $(T_m\Psi)^*$ is the transposed of the Jacobian matrix $\Psi'_{\rm loc}(\varphi^{-1}(m))$ and therefore $(\Psi^*\Xi)_{\rm loc}(x) = \Psi'_{\rm loc}(x)^T\,\Xi_{\rm loc}(x)$.

[2] In words: $\Psi_*\Theta$ acts on a vector field Y in the following way: Y is pulled-back to M with Ψ, its pull-back is paired with Θ in M, where Θ is defined, and the result is push-forward to N.

A.2 Lie derivatives and Lie brackets

We consider now the second nature of vector fields, that of derivations, even though we restrict this topic to the case we need—that of Lie derivatives of vector fields.

A.2.A Lie derivative of functions. We begin by recalling the notion and properties of Lie derivatives of functions. Even though elementary, this serves as a basis for the Lie derivative of vector fields.

Definition A.2.1 *The* Lie derivative *of functions associated to a vector field* $X \in \mathfrak{X}(M)$ *is the map* $L_X : C^\infty(M) \to C^\infty(M)$ *defined by*

$$L_X f = \langle df, X \rangle .$$

A *derivation* of an algebra $\mathcal{A}$ is a linear map $D : \mathcal{A} \to \mathcal{A}$ that satisfies Leibniz rule $D(f_1 f_2) = f_1 D f_2 + f_2 D f_1$ for all $f_1, f_2 \in \mathcal{A}$. Clearly, for any $X \in \mathfrak{X}(M)$, the Lie derivative L_X is a derivation of $C^\infty(M)$. It is immediate to verify that the space $Der(\mathcal{A})$ of all derivations of an algebra $\mathcal{A}$ is a vector space (over $\mathbb{R}$). We quote without proof the following fact:

Proposition A.2.2 *Let M be a manifold. The map $L : \mathfrak{X}(M) \to Der(C^\infty(M))$ given by $X \to L_X$ is a vector space isomorphism.*

Thus, every derivation of $C^\infty(M)$ is the Lie derivative associated to a vector field X su M.

There is a link between the flow of a vector field and the Lie derivative associated to it:

Proposition A.2.3 *For any $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$,*

$$L_X f = \tfrac{d}{dt}(f \circ \Phi_t^X)|_{t=0}$$

and, for any t,

$$\tfrac{d}{dt}(f \circ \Phi_t^X) = (L_X f) \circ \Phi_t^X .$$

Proof. Take $m \in M$. By the chain rule and (A.1.3), $\frac{d}{dt}(f \circ \Phi_t^X)(m)|_{t=0} = T_{\Phi_0^X(m)} f \cdot \frac{d}{dt}(\Phi_t^X)(m)|_{t=0} = T_m f \cdot X(m) = \langle df(m), X(m)\rangle_m = L_X f(m)$.

Fix now $\overline{t} \in \mathbb{R}$. Using (A.1.4), $\frac{d}{dt}(f \circ \Phi_t^X)(m)|_{t=\overline{t}} = \frac{d}{ds}(f \circ \Phi_{\overline{t}+s}^X(m)|_{s=0} = \frac{d}{ds}(f \circ \Phi_s^X \circ \Phi_{\overline{t}}^X(m)|_{s=0} = \frac{d}{ds}(f \circ \Phi_s^X)(\Phi_{\overline{t}}^X(m))|_{s=0} = (L_X f)(\Phi_{\overline{t}}^X(m))$. ■

Proposition A.2.4 *If $\Psi : M \to N$ is a diffeomorphism, then*

$$\Psi^*(L_Y g) \;=\; L_{\Psi^* Y} \Psi^* g \qquad \forall g \in C^\infty(N)\,, Y \in \mathfrak{X}(N)$$

An analogous identity for Ψ_* is also true.

Proof. Using (A.1.6) and Proposition A.2.3, $\Psi^*(L_Y g) = \Psi^*\langle dg, Y\rangle = \langle \Psi^* dg, \Psi^* Y\rangle = \langle d(\Psi^* g), \Psi^* Y\rangle = L_{\Psi^* Y}(\Psi^* g)$. ■

Exercises A.2.1 *(i) For vector fields $X^{\rm loc}$ and functions $f^{\rm loc}$ on open subsets of $\mathbb{R}^n$,*

$$L_{X^{\rm loc}} f^{\rm loc} = (f^{\rm loc})' X^{\rm loc} = X^{\rm loc} \cdot \nabla f^{\rm loc} = \sum_i X_i^{\rm loc} \frac{\partial f^{\rm loc}}{\partial x_i} .$$

(ii) In a local parametrization of M,

$$(L_X f)^{\rm loc} = L_{X^{\rm loc}} f^{\rm loc} .$$

(iii) Verify that L_X is a derivation.

(iv) Verify that, in a local parametrization φ with coordinates $(x_1, \ldots, x_n) = \varphi^{-1}$, the components $X_i^{\rm loc}$ of the representative $X^{\rm loc}$ of a vector field X are $X_i^{\rm loc} = (L_X x_i) \circ \varphi$.

(v) $X \in \mathfrak{X}(M)$ is the zero vector field if and only if $L_X f = 0$ for all $f \in C^\infty(M)$.

A.2.B Lie bracket and Lie derivative of vector fields. The Lie derivative and the Lie bracket (or commutator) of vector fields happen to be the same thing, but introduced in two different ways.

- The first way refers to the interpretation of vector fields as derivations of the algebra of functions. If $X, Y \in \mathfrak{X}(M)$, then the Lie derivatives $L_X, L_Y : C^\infty(M) \to C^\infty(M)$ are linear maps. Their commutator

$$L_X L_Y - L_Y L_X \ : \ C^\infty(M) \ \to \ C^\infty(M)$$

is a linear map which, in addition, satisfies Leibniz rule. This is verified observing that

$$\begin{aligned} L_X L_Y(fg) &= L_X(L_Y(fg)) \\ &= L_X(f L_Y g + g L_Y f) \\ &= (L_X f)(L_Y g) + f L_X(L_Y g) + (L_X g)(L_Y f) + g L_X(L_Y f) \end{aligned}$$

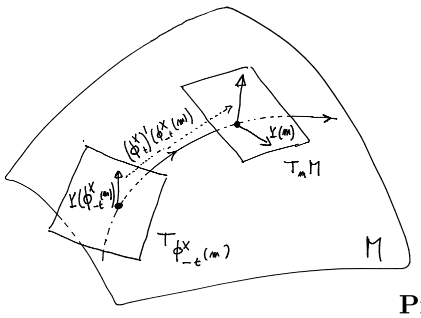
and that $L_Y L_X(fg)$ has the same expression with X and Y switched, so that

$$(L_X L_Y - L_Y L_X)(fg) = f(L_X L_Y - L_Y L_X)(g) + g(L_X L_Y - L_Y L_X)f .$$

This means that $L_X L_Y - L_Y L_X$ is a derivation of $C^\infty(M)$. Therefore, by Proposition A.2.2, there exists a vector field, which is denoted $[X, Y]$ and is called the *Lie bracket* or the *commutator* of X and Y, such that

$$L_{[X,Y]} f \ = \ (L_X L_Y - L_Y L_X) f \qquad \forall \ f \in C^\infty(M) . \tag{A.2.1}$$

- The second way refers to the idea of differentiating a vector field along the flow of another—similarly to what done for functions in Proposition A.2.3. Specifically, given two vector fields $X, Y \in \mathfrak{X}(M)$, the *Lie derivative of Y along X* is the vector field $L_X Y \in \mathfrak{X}(M)$ defined by



$$L_X Y(m) \;:=\; \tfrac{d}{dt}\big[(\Phi_t^X)^* Y\big](m)_{|t=0} \qquad \forall\, x \in M\,. \tag{A.2.2}$$

(That this actually defines a vector field on M is verified by noting that, for any $m \in M$, the r.h.s. is a vector in $T_m M$ because $[(\Phi_t^X)^* Y](m) = T_{\Phi_t^X(m)}\Phi_{-t}^X \cdot Y(\Phi_t^X(m))$ belongs to $T_m M$ for all t, and the derivative of a curve in a vector space is a vector of such a space).

Proposition A.2.5 $L_X Y = [X,Y]$ *for all* $X, Y \in \mathfrak{X}(M)$..

Proof. We may verify this equality in coordinates. To keep the notation simple, we we do not use different symbols for local representatives. Compute

$$\begin{aligned}
-L_X Y(x) &= \frac{d}{dt}\big[(\Phi_t^X)_* Y\big](x)\big|_{t=0} \\
&= \frac{d}{dt}\Big[\big[(\Phi_t^X)' Y \circ \Phi_{-t}^X\big](x)\Big]_{t=0} \\
&= \frac{d}{dt}\Big[(\Phi_t^X)'(\Phi_{-t}^X(x))\, Y(\Phi_{-t}^X(x))\Big]_{t=0} \\
&= \frac{d}{dt}\Big[(\Phi_t^X)'\big(\Phi_{-t}^X(x)\big)\Big]_{t=0} Y(x) + \frac{d}{dt}\Big[Y(\Phi_{-t}^X(x))\Big]_{t=0}
\end{aligned}$$

where we have used Leibniz rule, $\Phi_0^X = \mathrm{id}$, $(\Phi_0^X)'(x) = \mathbb{I}$ for all x. Now (keep in mind that $\Phi_t^X(x) = \Phi^X(t,x)$ is a function of the two variables t and x)

$$\frac{d}{dt}\Big[(\Phi_t^X)'\big(\Phi_{-t}^X(x)\big)\Big]_{t=0} \;=\; \Big[\frac{d(\Phi_t^X)'}{dt}(x)\Big]_{t=0} + (\Phi_0^X)''(x)\Big[\frac{d}{dt}\Phi_{-t}^X(x)\Big]_{t=0}.$$

Here the second term vanishes because $(\Phi_0^X)'(x) = \mathbb{I}$ is constant. Thus, using the 'variational equation' $\frac{d(\Phi_t^X)'}{dt}(x) = X'(\Phi_t^X(x))(\Phi_t^X)'(x)$ and again $(\Phi_0^X)'(x) = \mathbb{I}$, we have $\frac{d}{dt}\Big[(\Phi_t^X)'\big(\Phi_{-t}^X(x)\big)\Big]_{t=0} = X'(x)$. Since $\frac{d}{dt}\big[Y(\Phi_{-t}^X(x))\big]_{t=0} = Y'(x)\frac{d}{dt}\Phi_{-t}^X(x)\big|_{t=0} = -Y'(x)X(x)$, we eventually find $L_X Y(x) = -X'(x)Y(x) + Y'(x)X(x)$ which equals $[X,Y](x)$ (see Exercise A.2.2.i). ■

For this reason, the terms 'Lie derivatives' 'Lie bracket' and 'commutator' of vector fields are used as synonyms. We now see some of their properties:

Proposition A.2.6 *If M and N are manifolds and* $X, Y, Z \in \mathfrak{X}(M)$, *then:*

i. $[X,Y] = -[Y,X]$.

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

iii. $L_X(fY) = fL_XY + (L_Xf)Y$ *for all* $f \in C^\infty(M)$.

iv. $\frac{d}{dt}\big[(\Phi_t^X)^*Y\big] = (\Phi_t^X)^*L_XY$ *and* $\frac{d}{dt}\big[(\Phi_t^X)_*Y\big] = -(\Phi_t^X)_*L_XY$ *for all* $t \in \mathbb{R}$.

v. $\psi_*[X,Y] = [\psi_*X, \psi_*Y]$ *for any diffeomorphism* $\psi : M \to N$, *and similarly for the pull-back.*

Proof. i. and ii. follow, e.g., from the coordinate expression of the Lie bracket, which is given in Exercise A.2.2.i.

(iii.) In coordinates, $[(L_X(fY)]_i = L_X(fY_i) - L_{fY}X_i = Y_iL_Xf + fL_XY_i - fL_YX_i = Y_iL_Xf + f(L_XY)_i$.

(iv.) This can be verified with a computation similar to the one used to prove the analogous fact for the Lie derivative of functions, noticing also that $(\Phi_{t+s}^X)^* = (\Phi_s^X \circ \Phi_t^X)^* = (\Phi_t^X)^* \circ (\Phi_s^X)^*$.

(v.) This is equivalent to $L_{\psi_*[X,Y]}g = L_{[\psi_*X,\psi_*Y]}g$ for all functions $g \in C^\infty(N)$ (see Exercise A.2.1.v). Since Ψ is a diffeomorphism, any function on N is the push-forward of a function on M and we may thus equivalently verify that

$$L_{\psi_*[X,Y]}\psi_*f = L_{[\psi_*X,\psi_*Y]}\psi_*f \qquad \forall f \in C^\infty(M)\,.$$

Using the naturalness of the Lie derivatives of functions with respect to the push-forward we compute

$$\begin{aligned}
L_{[\psi_*X,\psi_*Y]}(\psi_*f) &= L_{\psi_*X}\big(L_{\psi_*Y}(\psi_*f)\big) - L_{\psi_*Y}\big(L_{\psi_*X}(\psi_*f)\big) \\
&= L_{\psi_*X}\big(\psi_*L_Yf\big) - L_{\psi_*Y}\big(\psi_*L_Xf\big) \\
&= \psi_*\big(L_XL_Yf - L_YL_Xf\big) \\
&= \psi_*(L_{[X,Y]}f) \\
&= L_{\psi_*[X,Y]}(\psi_*f)\,.
\end{aligned}$$

■

Statements i. and ii. of Proposition A.2.6 show that the Lie bracket is an anti-symmetric, bilinear map $\mathfrak{X}(M)\times\mathfrak{X}(M) \to \mathfrak{X}(M)$. The property in statement v. expresses the fact that the Lie bracket is natural with respect to push-forward and pull-back

We conclude recalling a basic result about the commutation of flows:

Proposition A.2.7 *For any pair of vector fields* X *and* Y *on a manifold, the following conditions are equivalent:*

i. $[X,Y] = 0$

ii. $(\Phi_s^Y)_*X = X$ *for all* $s \in \mathbb{R}$.

iii. The flows of X *and* Y *commute, namely* $\Phi_t^X \circ \Phi_s^Y = \Phi_s^Y \circ \Phi_t^Y$ *for all* $t, s \in \mathbb{R}$.

Proof. The equivalence of ii. and iii. follows from Proposition A.1.3 (with $\Psi = \Phi_s^Y$, $Z = (\Phi_s^Y)_*X$). That of i. and ii. from $\frac{d}{ds}(\Phi_s^Y)_*X = -(\Phi_s^Y)_*[X,Y]$ (and the fact that the push-forward under a diffeomorphism Ψ is injective, so that $\Psi_*Z = 0$ if and only if $Z = 0$). ■

Exercises A.2.2 *(i) Verify that the components* $[X,Y]_i^{\rm loc}$ *of the representative of* $[X,Y]$ *in a coordinate system are*

$$[X,Y]_i^{\rm loc} \;=\; L_{X^{\rm loc}} Y_i^{\rm loc} - L_{Y^{\rm loc}} X_i^{\rm loc} \;=\; (Y_i^{\rm loc})' X^{\rm loc} - (X_i^{\rm loc})' Y^{\rm loc}$$

and so

$$[X,Y]^{\rm loc} = (Y^{\rm loc})' X^{\rm loc} - (X^{\rm loc})' Y^{\rm loc} \, .$$

[Hint: If φ *is the local parametrization and* $\varphi^{-1} = (x_1, \ldots, x_n)$*, then* $[X,Y]_i^{\rm loc} = (L_{[X,Y]} x_i) \circ \varphi$ *(Exercise A.2.1.iii). Now use Exercises (A.2.1).i and ii.]*