Notes MA3407: Introduction to Lie theory

A. Schmeding

Draft Version 1.7

The present document is the draft version of notes prepared for the course

MA3407: Introduction to Lie theory

at NTNU in Fall 2024. These lecture notes present material from [HN12].[a]
If you find any errors please send a message to

alexander.schmeding@ntnu.no

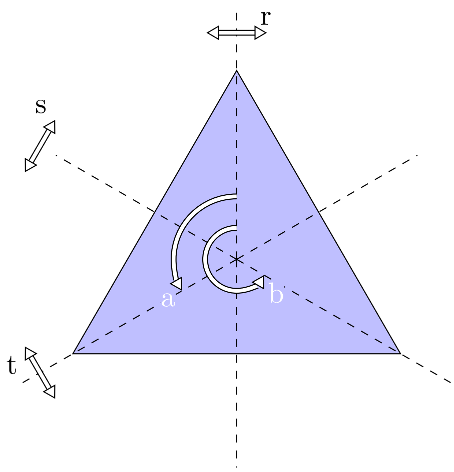
[a] For students at NTNU the book is available **for free** as a PDF via the library, see https://link.springer.com/book/10.1007/978-0-387-84794-8.

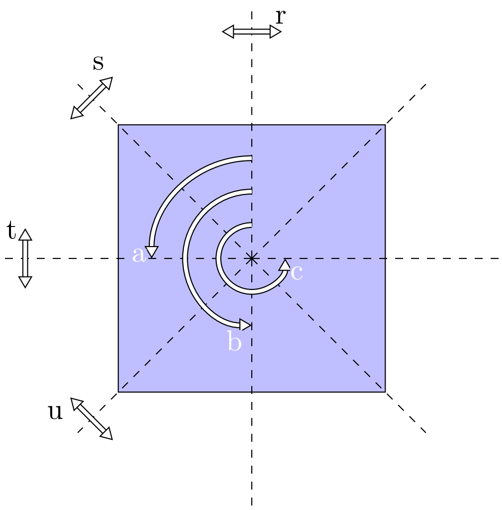
Contents

Introduction

In the first mathematics courses at university level, students meet analysis (including simple differential equations) and (linear) algebra (for example in the guise of vector spaces or as groups encoding symmetry). The topic of these notes sits at a meeting point for all of these ideas and is motivated by geometry. Lie theory deals with the structure and properties of Lie groups and Lie algebras. These are named after the Norwegian mathematician Marius Sophus Lie[1]. Sophus Lie investigated groups of symmetries associated to differential equations and called these groups "continuous groups" (a term which nowadays has been replaced with the term "Lie group"). The idea is that these objects can be studied using tools from algebra for the group level, ideas from topology ("continuity"!) and methods from analysis to gain insights into complicated systems such as differential equations.

To give an impression, of what this means, recall that regular objects like triangles or squares have finite groups of symmetries taking these objects to themselves:[2]





In general, there is no reason why interesting objects (for example in the plane) should have finite groups of symmetries taking them to themselves. For example the unit circle in R^2, i.e. all points (x, y) which satisfy the equation $x^2 + y^2 = 1$ has an infinite group

[1] 1842 (in Nordfjordeit) - 1899, Lie is one of the greatest Norwegian mathematicians who ever lived.
[2] Generated with TiKZ code by AndréCs answer to https://tex.stackexchange.com/questions/552589/square-rotational-and-reflection-symmetries.

of symmetries. For example, we can rotate the circle by any angle $\varepsilon \in \mathbb{R}$ which in polar coordinates $(x, y) = (\cos(\theta), \sin(\theta))$ pleasantly becomes the transformation

$$\Gamma : (x, y) \to (\cos(\theta + \varepsilon), \sin(\theta + \varepsilon)). \tag{0.1}$$

Observe that for fixed ε, the transformation Γ_ε has the following properties:

- Γ_ε is bijective and infinitely often differentiable in x, y, i.e. a smooth diffeomorphism of the plane.
- For $\varepsilon_1, \varepsilon_2$ we have $\Gamma_{\varepsilon_1} \circ \Gamma_{\varepsilon_2} = \Gamma_{\varepsilon_1+\varepsilon_2}$.
- the inverse of Γ_ε is $\Gamma_{-\varepsilon}$ and Γ_0 is the identity.

Summing up, a transformation is a symmetry, if it is a (smooth) diffeomorphism), preserves the structure and maps the object to itself. The symmetries form an infinite group of smooth transformations.

From (0.1) it is obvious that the transformation depends continuously (even infinitly often differentiably) on the parameter ε. We thus have a group of symmetries which is smoothly parametrised by ε, this structure is also called a 1-parameter Lie group.

First (naive) definition of 1-parameter groups To fix a bit of notation, assume that an object O is contained in $\mathbb{R}^n$ for some $n \in \mathbb{N}$ and has an infinite set of symmetries. We call a family of symmetries $\Gamma_\varepsilon, \varepsilon \in \mathbb{R}$ a 1-parameter group, if the following is satisfied

(a) $\Gamma_0 = \mathrm{id}_{\mathbb{R}^n}$ is the trivial symmetry.

(b) $\Gamma_\varepsilon \circ \Gamma_\delta = \Gamma_{\varepsilon+\delta}$ for every $\varepsilon, \delta \in \mathbb{R}$

(c) The map $\Gamma\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n, (x, \varepsilon) \to \Gamma_\varepsilon(x)$ is smooth (=infinitely often differentiable).

The fact that a 1-parameter symmetry group depends smoothly on some parameters provides a lot of additional information. In particular, these groups fit in nicely in the theory of differential equations. We will meet 1-parameter groups later again. In general, Lie groups are more complicated than 1-parameter groups. They typically form complicated non-linear objects called manifolds, but we shall first study easier examples of matrix Lie groups. Already for these it will turn out that the symmetry groups we are interested in have a linear object, the so called Lie algebra, attached to them. Lie theory investigates now the interplay between the Lie group and the associated Lie algebra. In these notes we shall give an introduction to basic principles and results which constitute what is nowadays called "Lie theory".

Lie symmetries of differential equations

We will briefly give an impression how the symmetry inherent in differential equations leads to 1-parameter groups. This was a classical motivation for Sophus Lie when investigating what later became known as a Lie group. The presentation here follows mostly [Sta07, Hyd00] and much more (quite practical) information can be found on this topic in these references.

In earlier courses we solved the first order ordinary differential equation (ODE)

$$y' = f(x) \tag{0.2}$$

by computing the following integrals

$$\int \mathrm{d}y = \int f(x)\mathrm{d}x \quad \Rightarrow \quad y = F(x) + c \text{ with } F' = f.$$

Another easy ODE is

$$y' = y \qquad \text{via} \quad \int \frac{\mathrm{d}y}{y} = \int \mathrm{d}x \quad \Rightarrow \quad \log(|y|) + c = x \Leftrightarrow y = ae^x \tag{0.3}$$

Indeed, both times the ODEs are *separable* i.e., we can separate terms involving only x from those involving only y.

A symmetry observation: The ODE (0.2) has solutions of the form $y = F(x) + c, c \in \mathbb{R}$, so solution curves are graphs of F moved in y-direction by the parameter c. Translating a solution in y-direction yields another solution. So the object we want to preserve when we talk about symmetry of the ODE is *the set of solution curves* for the ODE (recall that a unque solution is only selected after specifying an initial condition)! So a 1-parameter group of synnetries for the ODE (0.2) is given by

$$\varepsilon \mapsto \Gamma_\varepsilon(x, y) = (x, y + \varepsilon)$$

as it takes one solution curve $(x, F(x))$ in the plane to another one.

Similarly, for the ODE in (0.3), if we translate in x-direction the solution becomes $y = de^{x+c} = \hat{d}e^x$ for some parameter $\hat{d}$. So here translations in x-directions take a solution curve to a (different) solution curve. A 1-parameter curve of symmetries is thus given by

$$\varepsilon \mapsto \Gamma_\varepsilon(x, y) = (x + \varepsilon, y).$$

These examples are super simple and only toy examples as we can directly "see the symmetry". Note however, that symmetries are of interest precisely because they do not change the equation. Indeed, as they map solutions to solutions, they can be used to study general properties of the differential equations.

The idea is of course to figure out symmetries for difficult differential equations. Let us try with a different example which is not obvious. Consider the rational first order differential equation

$$\frac{dy}{dx} = \frac{y^3 + x^2 y - y - x}{xy^2 + x^3 + y - x} = \frac{y(y^2 + x^2 - 1) - x}{x(y^2 + x^2 - 1) + y} \tag{0.4}$$

It can be shown[3] that the equation transforms in polar coordinates $x = r\cos(\theta)$, $y = r\sin(\theta)$ to the much simpler separable equation

$$\frac{dr}{d\theta} = r(1 - r^2)$$

Again, translating now in the θ-direction (rotation in the usual coordinates!) takes a solution curve to another solution. A 1-parameter group of symmetries is thus

$$\varepsilon \to \Gamma_\varepsilon(x, y) = (\cos(\varepsilon)x + \sin(\varepsilon)y, -\sin(\varepsilon)x + \cos(\varepsilon)y).$$

This is far from obvious for the original equation, but again the trick was that secretly (0.4) was separable and separable differential equations have translation symmetries.

The obvious question is: When is it possible to find coordinates in which ODEs have symmetries? How does one find these transformations and what are the properties of the resulting symmetry groups? In these notes we shall only investigate the last of these questions. The others are treated in the sources mentioned in this section.

Literature on Lie theory

There are many books giving excellent introductions to the subject of Lie theory. Some introductory level texts are:

[3] Here is a physics style derivation: Lets pretend that differentials are some kind of fraction (they are not! But we are for a moment physicists and we use fractions instead of the chain rule...), so (0.4) measures the change with respect to time t of y divided by the change in x with respect to t (you know, fractions: $dy/dx = (dy/dt)/(dx/dt)$). Hence we solve the system

$$\begin{cases} \dot{y} = x(y^2 + x^2 - 1) + y \\ \dot{x} = x(y^2 + x^2 - 1) + y \end{cases}$$

Plug into the system now the identities $x = r\cos(\theta)$, $y = r\sin(\theta)$, $\dot{x} = \dot{r}\cos(\theta) - r\sin(\theta)\dot{\theta}$ and $\dot{y} = \dot{r}\sin(\theta) + r\cos(\theta)\dot{\theta}$. An unpleasant calculation later we find the system in matrix form becomes

$$\begin{pmatrix} \dot{r} \\ \dot{\theta} \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} r(1 - r^2) \\ 1 \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

So $\dot{r} = r(1 - r^2)$ and fortunately, $d\theta/dt = \dot{\theta} = 1$. Smuggling in $\dot{\theta} = d\theta/dt$ into $\dot{r} = dr/dt$ (this is of course the chain rule applied to get rid of our fraction trick, exploiting that the correction factor is 1!) we get $\frac{dr}{d\theta} = r(1 - r^2)$

- Duistermaat, J.J. and Kolk, A.: Lie groups, [DK00], download via Springer-link: https://link.springer.com/book/10.1007/978-3-642-56936-4
- Hilgert/Neeb [HN12], download via Springer link: https://link.springer.com/book/10.1007/978-0-387-84794-8.
- Hall's book [Hal15] can be downloaded via Springer link: https://link.springer.com/book/10.1007/978-3-319-13467-3

Somewhat more advanced (and with connections to a variety of topics):

- S. Helgason: Differential geometry, Lie groups, and symmetric spaces. [Hel01]
- J.E. Humphreys: Introduction to Lie algebras and representation theory. 3rd printing, rev, [Hum80]
- A.W. Knapp: Lie groups, beyond an introduction [Kna02]
- General Lie groups and infinite-dimensional Lie groups, [Sch23].

Applications of Lie theory

Lie theory got its start when mathematicians tried to figure out ways to use symmetry to solve differential equations. Here are some pointers to the literature:

- A nice introductory overview is in the article by Starret, [Sta07].
- A guide for applied mathematicians and physicists on how to apply symmetry methods is found in Hydon's book, [Hyd00].

More advanced in this direction is the book:

- P. Olver: Applications of Lie Groups to Differential Equations, [Olv93].

Conventions and notation

Throughout these notes we write $\mathsf{N} = \{1, 2, 3, \ldots\}$ for the *natural numbers* and $\mathsf{N}_0 = \mathsf{N} \ \ \{0\}$. Further K will always denote either the *field of real numbers* R or the *field of complex numbers* C.
We need a bit of topology (see Appendix A.1 for some basic notions) though mostly this will mean that we work in metric spaces. However, we shall write $U \ {}^{\circ}\ X$ if we want to say that U is an open subset of a topological space X. Whenever we consider the cartesian product $X \times Y$ of two topological spaces X, Y, we will endow it with the product topology if nothing else is said.

1. Matrix Lie groups

In this chapter we will develop the basic Lie theory of matrix Lie groups. For this we will work on a finite dimensional vector space and construct groups of matrices. The group operation will always be matrix multiplication. We shall refrain from a formal definition of the term "matrix Lie group" (informally we will mostly equate them with linear Lie groups which are closed subgroups of the general linear group we discuss in the next section). Our aim is to see that matrix Lie groups are to a surprising amount determined by a (on first sight) much simpler structure, their Lie algebra.

Conventions: Matrices and norms on finite dimensional vector spaces

We consider $\mathbb{K}^n$ as a normed space with respect to the euclidean norm

$$\|x\| = \sqrt{\langle x, x\rangle}, \qquad \text{where } \langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y}_i,$$

is the standard inner product with $\overline{y}$ the complex conjugate of y.
Write $M_n(\mathbb{K})$ for the set of all $n \times n$-matrices with entries in $\mathbb{K}$. Whenever we work with a matrix A we use A_{ij} as a shorthand for the entry in the ith row and jth column. Now endow $M_n(\mathbb{K})$ with the operator norm with respect to $\|\cdot\|$:

$$\|A\| := \sup_{\|v\|=1} \|Av\|.$$

Note $(M_n(\mathbb{K}), \|\cdot\|)$ is a normed space[a] and the matrix product then satisfies $\|A \cdot B\| \leq \|A\| \cdot \|B\|$.

[a] Looking at the entries of the matrices we can identify it with $\mathbb{K}^{n^2}$.

1.1. The general linear group

1.1.1 Definition Let $\mathrm{GL}_n(\mathbb{K})$ be the group of all invertible $n \times n$-matrices (the group structure is given by the matrix product). The unit of the group is the identity matrix I_n.

1.1.2 Proposition *The group* $\mathrm{GL}_n(\mathbb{K})$ *is an open subset of* $M_n(\mathbb{K})$. *Both the product and inversion maps are smooth and thus in particular continuous.*

The classical proof [HN12] of Proposition 1.1.2 uses the Leibniz formula for the determinant and Cramer's rule for the inverse. These are NOT taught in Matematikk 3 at NTNU[1] We shall follow a different route here.

Proof of Proposition 1.1.2. **Step 1:** *The matrix product* $m: M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to M_n(\mathbb{K})$ *is smooth*: For matrices $A = (A_{ij}), B = (B_{ij})$ the product is given by the matrix

$$(A \cdot B)_{ij} = \left(\sum_{k=1}^{n} A_{ik} B_{kj}\right)_{ij}$$

So every component of $m(A, B)$ is a polynomial of the entries of the matrices A and B. Polynomials are smooth and since $M_n(\mathbb{K}) = \mathbb{K}^{n^2}$ smoothness of the product is equivalent to smoothness of the component maps. In conclusion, the product is smooth. Let us compute its directional derivative at a pair of matrices X, Y:

$$\begin{aligned} \mathrm{d}m(X, Y)(H_1, H_2) &= \lim_{t\to 0} t^{-1}(m(X + tH_1, Y + tH_2) - m(X, Y)) \\ &= \lim_{t\to 0} t^{-1}((X + tH_1) \cdot (Y + tH_2) - XY) \\ &= \lim_{t\to 0} t^{-1}(t(XH_2 + H_1Y) + t^2 H_1 H_2) = XH_2 + H_1 Y \qquad (1.1) \end{aligned}$$

Step 2: *Inversion of matrices is smooth near the identity matrix* I_n: Apply the implicit function theorem to $F(X, Y) := m(X, Y) - I_n$. For this we note:

- F is smooth (with the same Jacobian as m).
- $F(I_n, I_n) = 0$
- Identifying $M_n(\mathbb{K}) = \mathbb{K}^{n^2}$, the Jacobian matrix $J_F(X, Y)$ is a $n^2 \times 2n^2$ matrix. For $X = I_n = Y$, (1.1) yields $J_F(I_n, I_n) = \begin{pmatrix} I_{n^2} & I_{n^2} \end{pmatrix}$.

Hence the Implicit function Theorem A.2.8 shows that there is an open I_n-neighborhood U together with a smooth function $\iota: U \to M_n(\mathbb{K})$ such that $F(X, \iota(X)) = X \cdot \iota(X) - I_n = 0$. By uniqueness of the inverse $\iota(X) = X^{-1}$ and inversion is smooth near the identity.

Step 3: $\mathrm{GL}_n(\mathbb{K})$ *is open in* $M_n(\mathbb{K})$: As $\mathrm{GL}_n(\mathbb{K}) = \det^{-1}(\mathbb{K} \setminus \{0\})$ and $\mathbb{K} \setminus \{0\}$ is an open set, the claim follows as the determinant is a continuous map (indeed even smooth, cf. Exercise 1.1). (preimages of open sets under a continuous maps are open).

Step 4: *The group operations are smooth on* $\mathrm{GL}_n(\mathbb{K})$: Multiplication is smooth by Step 1 as smooth maps restrict to smooth maps on the open set $\mathrm{GL}_n(\mathbb{K})$. By step 2, inversion

[1] but check Wikipedia https://en.wikipedia.org/wiki/Leibniz_formula_for_determinants and https://no.wikipedia.org/wiki/Cramers_regel.

is smooth on the open identity neighborhood U. Let $A \in \mathrm{GL}_n(\mathbb{K})$ be arbitrary. Since $A^{-1} \cdot A = \mathrm{id}$, there is an open neighborhood V_A such that multiplication with A^{-1} maps V_A into U. We prove that on V_A inversion is smooth. By Step 2

$$Z^{-1} = (A^{-1} \cdot Z)^{-1} \cdot A^{-1} = m(\iota(m(A^{-1}, Z)), A^{-1}), \qquad Z \in V_A$$

is a smooth as composition of the smooth maps m and ι. This concludes the proof. □

Another way of formulating parts of Proposition 1.1.2, is to say that the group $\mathrm{GL}_n(\mathbb{K})$ is a topological group.

1.1.3 Definition Let G be a group which is also a topological space. We call G a *topological group* if group multiplication $m \colon G \times G \to G$ and inversion $\iota \colon G \to G$ are continuous.

See Appendix B for more information on topological groups. Indeed we already proved that $\mathrm{GL}_n(\mathbb{K})$ is a group with smooth group operations. Here we used that the group is an open subset of a vector space to make sense of smoothness. Open subsets of vector spaces are manifolds, whence the following definition is natural.

1.1.4 Definition Let G be a group which is also a manifold[2] such that multiplication $m \colon G \times G \to G$ and inversion $\iota \colon G \to G$ are smooth. Then we call G a *Lie group*.

1.1.5 Example
- The group $\mathrm{GL}_n(\mathbb{K})$ is a Lie group.
- The invertible complex numbers $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$ with multiplication form a Lie group. In Exercise 1.1.1 we will identify this group with a matrix Lie group,

We will now collect more examples of these groups with additional structure.

Conventions: transpose and complex conjugate of matrices

For $A \in M_n(\mathbb{K})$ we will now (and in the following) write $A = (A_{ij})_{1 \le i,j \le n}$ also as (A_{ij}) for the components. Then we recall

$$A^\top = (A_{ji}), \qquad \overline{A} = (\overline{A}_{ij}), \qquad A^* = \overline{A}^\top = (\overline{A}_{ji})$$

Recall that $(A \cdot B)^\top = B^\top \cdot A^\top$. Further, $A^\top = A^*$ is equivalent to $A = \overline{A}$ (whence all entries of such a matrix need to be real numbers).

[2] In case you do not know what a manifold is: A manifold is a set on which we can make sense of differentiable maps. We will return to this in a later chapter. For the time being note that open subsets of vector spaces are manifolds.

We can now introduce notation for some important subgroups of $\mathrm{GL}_n(\mathsf{K})$:

1.1.6 Definition We define the following subgroups of $\mathrm{GL}_n(\mathsf{K})$:

- The *special linear group* $\mathrm{SL}_n(\mathsf{K}) \coloneqq \{A \quad \mathrm{GL}_n(\mathsf{K}) \mid \det(A) = 1\}$.
- The *orthogonal group* $\mathrm{O}_n(\mathsf{K}) \coloneqq \{A \quad \mathrm{GL}_n(\mathsf{K}) \mid A \quad = A^{-1}\}$.
- The *special orthogonal group* $\mathrm{SO}_n(\mathsf{K}) \coloneqq \mathrm{SL}_n(\mathsf{K}) \cap \mathrm{O}_n(\mathsf{K})$
- The *unitary group* $\mathrm{U}_n(\mathsf{K}) = \{A \quad \mathrm{GL}_n(\mathsf{K}) \mid A \quad = A^{-1}\}$. Note that $\mathrm{U}_n(\mathsf{R}) = \mathrm{O}_n(\mathsf{R})$ but $\mathrm{U}_n(\mathsf{C}) = \mathrm{O}_n(\mathsf{C})$.
- The *special unitary group* $\mathrm{SU}_n(\mathsf{K}) \coloneqq \mathrm{SL}_n(\mathsf{K}) \cap \mathrm{U}_n(\mathsf{K})$.

We (this means you) will check in Exercise 1.1 2. below that these sets are indeed subgroups of the general linear group.

1.1.7 Remark Every subgroup of a topological group (with the subspace topology) is a topological group, Lemma B.0.1. So the subgroups of $\mathrm{GL}_n(\mathsf{K})$ from Definition 1.1.6 are topological groups. Moreover, we shall see (for example in Lemma 1.2.1) below that all of the groups in Definition 1.1.6 are closed subgroups.

All of the subgroups of $\mathrm{GL}_n(\mathsf{K})$ we defined so far are closed subgroups. In a certain sense this is no accident, at a bare minimum it rules out a lot of undesirable examples such as the following:

1.1.8 Example Let $\mathrm{GL}_n(\mathsf{Q})$ be the set of invertible $n \times n$-matrices with rational coefficients. These matrices form a subgroup of $\mathrm{GL}_n(\mathsf{K})$ (using the formula for the inverse via the adjugate matrix one sees that the inverse has again rational coefficients), but the subgroup is not closed. In particular, we stated in the introduction that we are potentially interested in 1-parameter groups, i.e. groups which depend smoothly on a parameter ε. These are quite boring for the subgroup $\mathrm{GL}_n(\mathsf{Q})$ as they are all constant.

Focusing on closed subgroups of matrix groups we can avoid many pathological examples which do not fit in the theory.

We said that a Lie group is supposed to be a manifold. We will see later in this course that closed subgroups of $\mathrm{GL}_n(\mathsf{K})$ admit a compatible manifold structure turning them into Lie groups in the sense of Definition 1.1.4 (this is a deep theorem and not obvious at all). Let us ignore for a moment that we have not yet introduced the concept of a manifold and cheat:

1.1.9 Definition Let $U \subseteq \mathbb{R}^n$ and $A \subseteq U$ some subset. We say that a map $g\colon A \to \mathbb{R}^m$ is a *C^k-map*, $k \in \mathbb{N}_0 \cup \{\infty\}$ if there exists a C^k-map $\hat{g}\colon U \to \mathbb{R}^m$ such that its restriction to A satisfies $\hat{g}|_A = g$.

So since the group operations of $\mathrm{GL}_n(\mathbb{K})$ are smooth we obtain smooth group operations on the subgroups from Definition 1.1.6. We will see later that all of these groups are submanifolds of $\mathrm{GL}_n(\mathbb{K}) = \mathbb{R}^{2n}$ and for closed subgroups, smoothness in the sense of Definition 1.1.9 coincides with smoothness with respect to the (sub)manifold structure. We give this special situation a name:

1.1.10 Definition A closed subgroup of $\mathrm{GL}_n(\mathbb{K})$ is called a *linear Lie group*.

In the first part of this lecture we shall exclusively study linear Lie groups (colloquially also often called matrix Lie groups).

1.1.11 Corollary *The Lie group $\mathrm{GL}_n(\mathbb{K})$ and the subgroups from Definition 1.1.6 are linear Lie groups.*

1.1.12 Example (Finite groups) A matrix $P \in M_n(\mathbb{K})$, $n \in \mathbb{N}$ is called permutation matrix if it contains in every row and column exactly once the entry 1 and otherwise only 0. Denote by

$$S_n := \{P \in M_n(\mathbb{K}) \mid P \text{ is a permutation matrix}\}.$$

Then S_n is a finite subgroup of $\mathrm{O}_n(\mathbb{R})$ and by finiteness, S_n is in particular closed in $\mathrm{GL}_n(\mathbb{K})$ whence a linear Lie group.

As an abstract group S_n is isomorphic to the group $\mathrm{Sym}(n)$ of all permutations of a set with n elements. Now Cayley's theorem (see e.g. [Jac85]) asserts that every finite group is isomorphic to a subgroup of $\mathrm{Sym}(n)$ for some n. Thus every finite group can be interpreted as a (linear) Lie group.

1.1.13 Example (Heisenberg group) Consider the subgroup

$$H := \left\{ \begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} \middle| \, x, y, z \in \mathbb{R} \right\}$$

of $\mathrm{GL}_3(\mathbb{R})$. The group H is called *Heisenberg group*. It is clearly closed in $\mathrm{GL}_3(\mathbb{R})$ and thus a linear Lie group. The Heisenberg group occurs in a variety of applications from quantum mechanics, harmonic analysis to sub-Riemannian geometry.

1.1 Exercises

1. Identify a complex number $z = x + iy$ with the 2×2 real matrix $\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$

a) Show that the identification takes multiplication in $\mathbb{C}$ to matrix multiplication.

b) Let $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Show that matrices $M \in M_2(\mathbb{R})$ that are of the above form are exactly those such that $JM = MJ$. (note that J represents i).

c) Show that the group $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$ of invertible complex numbers (wrt. multiplication) corresponds to a closed subgroup of $\mathrm{GL}_2(\mathbb{R})$.

2. Show that $\mathbb{R}$ with addition can be identified as a closed subgroup of $M_2(\mathbb{R})$.
Hint: Consider the matrices $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ for $\lambda \in \mathbb{R}$

3. We will prove in this exercise that the determinant $\det\colon M_n(\mathbb{K}) \to \mathbb{K}$ is a polynomial in the matrix entries, whence it is smooth and continuous for every $n \in \mathbb{N}$. Recall that $M_n(\mathbb{K}) = \mathbb{K}^{n^2}$.

a) Deduce from the identification that for $1 \leq i, j \leq n$ the map $M_n(\mathbb{K}) \to \mathbb{K}, A \to A_{ij}$ sending a matrix to the entry A_{ij} is smooth.

b) Explain why for each $1 \leq k, \ell \leq n$ the map $\delta_{k,\ell} : M_n(\mathbb{K}) \to M_{n-1}(\mathbb{K})$ sending a matrix to the matrix where we deleted the kth row and the ℓth column is smooth.

c) Use induction and the Laplace formula for the determinant

$$\det A = \sum_{i=1}^{n} (-1)^{1+i} A_{1i} \det(\delta_{1,i}(A))$$

to prove that the determinant is a polynomial in the entries of the matrix.
Hint: Sums and products of polynomials are polynomials!

4. Use the formula $\det(A \cdot B) = \det(A) \cdot \det(B)$ to prove that the sets in Definition 1.1.6 indeed form subgroups of $\mathrm{GL}_n(\mathbb{K})$.

5. Let $G_i, 1 \leq i \leq n$ be linear Lie groups and consider the direct product group $G = G_1 \times G_2 \times \cdots \times G_n$ (cf. Exercise B.1) Show that the direct product is again a linear Lie group.
Hint: Can you identify the direct product as a block diagonal matrix?

1.2. Topological structure of matrix Lie groups

The matrix groups we considered so far showcase that Lie groups blend algebraic and analytic theory. For later it will be useful to have some knowledge about the topological structure of these matrix groups.

There are two properties which will be relevant for us: compactness and (path) connectedness.[3] Compactness is a smallness condition. Let us note first that $\mathrm{GL}_n(\mathsf{K})$ is not compact (exercise). However, many of the subgroups we introduced are compact.

1.2.1 Lemma *For every $n \quad \mathsf{N}$, the groups*

$$\mathrm{U}_n(\mathsf{C}), \quad \mathrm{SU}_n(\mathsf{C}), \quad \mathrm{O}_n(\mathsf{R}), \text{ and } \mathrm{SO}_n(\mathsf{R})$$

are compact.

Proof. We identify $M_n(\mathsf{C}) = \mathsf{C}^{n^2}$ and have to show that the subsets are closed and bounded (employing the Heine-Borel Theorem).

Boundedness: In view of

$$\mathrm{SO}_n(\mathsf{R}) \quad \mathrm{O}_n(\mathsf{R}) \quad \mathrm{U}_n(\mathsf{C}), \text{ and } \mathrm{SU}_n(\mathsf{C}) \quad \mathrm{U}_n(\mathsf{C}),$$

it suffices to prove that $\mathrm{U}_n(\mathsf{C})$ is bounded. If $g_1 \cdots g_n$ are the columns of the matrix G, then $G \; = G^{-1}$ is equivalent to $GG \; = \mathrm{id}_n$, which means that $g_1, \ldots, g_n$ form an orthonormal basis for C^n with respect to the inner product $\; x, y \; = \; {}_i x_i \bar{y}_i$ which induces the norm via $\; z \; = \; \overline{z, z} \;$. Thus $\; g_i \; = 1$ for each i if $G \quad \mathrm{U}_n(\mathsf{C})$ and thus the set $\mathrm{U}_n(\mathsf{C})$ is bounded.

Closed The functions

$$f, h \colon M_n(\mathsf{K}) \to M_n(\mathsf{K}), f(A) = AA \; - \mathrm{id}_n \text{ and } h(A) = AA \; - \mathrm{id}_N$$

are continuous. Therefore, the groups

$$\mathrm{U}_n(\mathsf{K}) = f^{-1}(0) \text{ and } \mathrm{O}_n(\mathsf{K}) = h^{-1}(0)$$

are closed. Likewise $\mathrm{SL}_n(\mathsf{K}) = \det^{-1}(1)$ is closed and therefore the group $\mathrm{SU}_n(\mathsf{C})$ and $\mathrm{SO}_n(\mathsf{R})$ are closed as the intersection of two closed subsets. □

We now turn to path connectedness which means that two points can be connected using a continuous path.[4] So it is a natural property we would assume is true at least locally (1-parameter groups certainly are path connected!).

[3] See Appendix A.1 for the defintions in case you are not sure.

[4] In a topology course one learns about the property of a topological space X being connected. This means that if $X = U \quad V$ and both U, V are open with $U \cap V = \;$, then either $U = X$ or $U = \;$. Path connectedness is a stronger notion than connectedness. However, since we work on manifolds (which locally look like euclidean space) the two notions actually coincide.

1.2.2 Proposition *The groups* $\mathrm{GL}_n(\mathbb{C})$*,* $U_n(\mathbb{C})$ *and* $\mathrm{SL}_n(\mathbb{C})$ *are path connected for all* $n \in \mathbb{N}$*.*

Proof. We start with $\mathrm{GL}_n(\mathbb{C})$: Let $X, Y \in \mathrm{GL}_n(\mathbb{C})$, then the convex combination $c\colon [0,1] \to M_n(\mathbb{C}), c(t) = tX + (1-t)Y$ is a continuous path starting at Y and ending at X. Unfortunately, there is no reason its image lies in $\mathrm{GL}_n(\mathbb{C})$. However, since the determinant is a polynomial in the entries of the matrix, $\mathbb{C} \to \mathbb{C}, z \to \det(zX + (1-z)Y)$ is a complex polynomial of degree n. By the fundamental theorem of algebra it has exactly n-zeroes and we can thus pick a path $p\colon [0,1] \to \mathbb{C}$ with $p(0) = 0, p(1) = 1$ that avoids the zeroes of this polynomial. Then $\tilde{c}(t) := p(t)X + (1 - p(t))Y$ is a path from Y to X in $\mathrm{GL}_n(\mathbb{C})$. This shows that $\mathrm{GL}_n(\mathbb{C})$ is path connected.

Now if $X, Y \in \mathrm{SL}_n(\mathbb{C})$ consider instead $\hat{c}(t) = (\det \tilde{c}(t))^{-1/n}\tilde{c}(t)$ to get a path in $\mathrm{SL}_n(\mathbb{C})$ (where we have to possibly adjust choices of p so that it does not wind around 0 and we can use a continuous branch of the nth root).

Finally, for $U_n(\mathbb{C})$: Pick $U \in U_n(\mathbb{C})$ From linear algebra we know that there exists an orthonormal base $v_1, \dots, v_n$ of eigenvectors of U. Let $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues. Since U is unitary, $|\lambda_j| = 1$, whence there is $\theta_j \in \mathbb{R}$ with $\lambda_j = e^{\theta_j i}$. Now we define the path

$$\gamma\colon [0,1] \to \mathrm{U}_n(\mathbb{C}), \quad \gamma(t)v_j := e^{t\theta_j i}v_j, j = 1, \dots, n.$$

By construction $\gamma(0) = I_n$ and $\gamma(1) = U$. (for all other values of t, $\gamma(t)$ is a unitary matrix as the v_j form an orthonormal basis). □

1.2.3 Example The groups $\mathrm{GL}_n(\mathbb{R})$ and $\mathrm{O}_n(\mathbb{R})$ are not path connected for any $n \in \mathbb{N}$, To see this, note that as det is continuous, $\mathrm{GL}_n^-(\mathbb{R}) = \det^{-1}(]-\infty, 0[)$ and $\mathrm{GL}_n^+(\mathbb{R}) = \det^{-1}(]0, \infty[)$ are open disjoint subsets of $\mathrm{GL}_n(\mathbb{R})$ which partition the set into two open (and closed) pieces. Hence $\mathrm{GL}_n(\mathbb{R})$ is not connected and thus not path connected. The same argument applies to $\mathrm{O}_n(\mathbb{R})$.

1.2.4 Lemma *The orthogonal group* $\mathrm{O}_n(\mathbb{R})$ *consists of two path components*[5]

$$\mathrm{SO}_n(\mathbb{R}) \text{ and } \mathrm{O}_n^-(\mathbb{R}) = \mathrm{O}_n(\mathbb{R}) \cap \mathrm{GL}_n^-(\mathbb{R}).$$

Proof. You will show in Exercise 1.2.3 that $\mathrm{SO}_n(\mathbb{R})$ is path connected. As a subgroup it contains the unit and $\mathrm{SO}_n(\mathbb{R}) = \mathrm{O}_n(\mathbb{R}) \cap \mathrm{GL}_n^+(\mathbb{R})$. From the discussion in Example 1.2.3 we can guess that $\mathrm{O}_n^-(\mathbb{R})$ is the other path component if it is path connected. Let us prove this now: If $A, B \in \mathrm{O}_n^-(\mathbb{R})$, then $A^{-1}B \in \mathrm{SO}_n(\mathbb{R})$ and we can pick a path $\gamma\colon [0,1] \to \mathrm{SO}_n(\mathbb{R})$ connecting I_n and $A^{-1}B$. But then $p(t) = A\gamma(t)$ is a path in $\mathrm{O}_n^{-1}(\mathbb{R})$ connecting $A\gamma(0) = AI_n = A$ with $A\gamma(1) = A(A^{-1}B) = B$. So $\mathrm{O}_n^-(\mathbb{R})$ is a path component of $\mathrm{O}_n(\mathbb{R})$. □

[5] For the definition of path components see Definition A.1.9

1.2 Exercises

1. Show that the following groups are not compact:

a) $\mathrm{GL}_n(\mathbb{K})$ for arbitrary $n \in \mathbb{N}$

b) $\mathrm{SL}_n(\mathbb{K})$ for $n > 1$ (what happens for $n = 1$?)
Hint: Show that the the groups do not satisfy the conditions in the Heine-Borel theorem.

2. We investigate path connectedness of $\mathrm{GL}_n(\mathbb{K})$:

a) Find examples of matrices $A, B \in \mathrm{GL}_n(\mathbb{C})$ such that the path $c(t) = tA + (1-t)B$ is NOT completely contained in $\mathrm{GL}_n(\mathbb{C})$.

b) For the proof of Proposition 1.2.2: Draw a picture which explains why we can always choose the path p in $\mathbb{C}$ to modify $c(t) = tA + (1-t)B$ to a path in $\mathrm{GL}_n(\mathbb{C})$. In particular: Why is it allowed to pick one with $p(0) = 0$ and $p(1) = 1$? Then explain why this does not work for $\mathrm{GL}_n(\mathbb{R})$.

c) Show that the path component of the unit in $\mathrm{GL}_n(\mathbb{R})$ is contained in the matrices with determinant > 0 (indeed one can show that they coincide)

3. Prove by induction on n that the groups $\mathrm{SO}_n(\mathbb{R})$ are path connected. For this:

a) Argue that for $n = 1$ there is nothing to do, so we may assume $n \geq 2$.

b) Write e_1 for the first standard unit vector in $\mathbb{R}^n$. Show that for every unit vector $v \in \mathbb{R}^n$ there exists a continuous path $r\colon [0,1] \to \mathrm{SO}_n(\mathbb{R})$ with $r(0) = I_n$ and $r(1)v = e_1$.

c) Show that any $X \in \mathrm{SO}_n(\mathbb{R})$ can be connected by a continuous path to a block-diagonal matrix of the form $\begin{pmatrix} 1 & 0 \\ 0 & X_{n-1} \end{pmatrix}$ with $X_{n-1} \in \mathrm{SO}_{n-1}(\mathbb{R})$ and complete the proof by induction.

1.3. Groups and geometry

In Definition 1.1.6 we defined closed subgroups of $\mathrm{GL}_n(\mathbb{K})$ by imposing certain condition on the matrices. It is often better to think of matrices as linear maps described with respect to some basis and thus to adopt a more abstract point of view. The point is that the groups we investigated earlier all appear as symmetry groups of certain bilinear forms on a finite dimensional vector space. This connection to geometry is worth spelling out as it does not require us to fix a base of the vector space

1.3.1 Definition Let V be a $\mathbb{K}$-vector space. We write $\mathrm{GL}(V)$ for the group of linear automorphisms of V (i.e. all invertible elements of the space of linear maps $\mathrm{Lin}(V, V)$).

If V is n-dimensional and $v_1, \ldots, v_n$ a basis, then the map

$$\Phi\colon M_n(\mathbb{K}) \to \mathrm{Lin}(V, V), \quad \Phi(A)v_k := \sum_{j=1}^{n} A_{jk} v_j$$

is a linear isomorphism which describes the passage between linear maps and matrices. Note that since $\Phi(I_n) = \mathrm{id}_V$ and $\Phi(AB) = \Phi(A)\Phi(B)$, Φ restricts to a group isomorphism

$$\Phi_{\mathrm{GL}_n(\mathbb{K})}\colon \mathrm{GL}_n(\mathbb{K}) \to \mathrm{GL}(V).$$

For a *bilinear map* $\beta\colon V \times V \to \mathbb{K}$ we use our choice of basis to define the matrix $B = (B_{jk}) := (\beta(v_j, v_k))_{j,k=1,\ldots,n}$. If we use the basis of V to identify V with $\mathbb{K}^n$ (by using the coordinates induced by the basis), the matrix B describes β as a map on $\mathbb{K}^n$ via $\beta(x, y) = x^\top By$. Thus B behaves differently from the matrix of a linear map!

Write

$$\mathrm{Aut}(V, \beta) := \{g \in \mathrm{GL}(V)\colon (\forall v, w \in V)\ \beta(gv, gw) = \beta(v, w)\}$$

for the *isometry group of the bilinear form* β. Then clearly

$$\Phi^{-1}(\mathrm{Aut}(V, \beta)) = \{g \in \mathrm{GL}_n(\mathbb{K})\colon g^\top Bg = B\}$$

If β is *symmetric*, i.e. $\beta(v, w) = \beta(w, v)$ for all $v, w \in V$, we also write $\mathrm{O}(V, \beta) := \mathrm{Aut}(V, \beta)$ and if β is *skew-symmetric*, i.e. $\beta(v, w) = -\beta(w, v)$ for all $v, w \in W$, we write $\mathrm{Sp}(V, \beta) := \mathrm{Aut}(V, \beta)$.

If $v_1, \ldots, v_n$ is an orthonormal basis[6] for β, i.e. the associated matrix $B = I_n$, then

$$\Phi^{-1}(\mathrm{Aut}(V, \beta)) = \mathrm{O}_n(\mathbb{K}), \tag{1.2}$$

is the orthogonal group from Definition 1.1.6. Note that orthonormal bases can only exist for symmetric bilinear forms (Why?). So (1.2) yields $\Phi^{-1}(\mathrm{O}(V, \beta)) = \mathrm{O}_n(\mathbb{K})$.

For $V = \mathbb{K}^{2n}$ and the block $(2n \times 2n)$-matrix

$$B = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

we have $B^\top = -B$, and the group

$$\mathrm{Sp}_{2n}(\mathbb{K}) := \{g \in \mathrm{GL}_n(\mathbb{K})\colon g^\top Bg = B\}$$

[6] Due to the formula for the associated matrix an orthonormal basis is one where $\beta(v_i, v_j)$ equals 1 if $i = j$ and 0 otherwise. This justifies the terminology.

is called the ***symplectic group***. This group is considered in symplectic geometry. The corresponding skew-syemmtric bilinear form on $\mathbb{K}^{2n}$ is given by

$$\beta(x,y) = x^\top By = \sum_{i=1}^{n} x_i y_{n+i} - x_{n+i} y_i.$$

1.3.2 Definition A symmetric bilinear form β on V is called ***non degenerate*** if $\beta(v,V) := \{\beta(v,w)\colon w \in V\} = \{0\}$ implies $v = 0$.

For $\mathbb{K} = \mathbb{C}$ every non degenerate symmetric bilinear form β possesses an orthonormal basis, so that for every such β we have an isomorphism of groups

$$\mathrm{O}(V,\beta) = \mathrm{O}_n(\mathbb{C}).$$

For $\mathbb{K} = \mathbb{R}$ the situation is more complicated since negative real numbers do not posses (real) roots. However one can show that every non degenerate bilinear form on an $\mathbb{R}$-vector space admits an orthogonal basis such that the matrix associated to β can be written as

$$I_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}, \text{ for some } p+q=n.$$

Then $O(V,\beta)$ is isomorphic to the group

$$\mathrm{O}_{p,q}(\mathbb{R}) = \{g \in \mathrm{GL}_n(\mathbb{R})\colon g^\top I_{p,q} g = I_{p,q}\},$$

where $\mathrm{O}_{n,0}(\mathbb{R}) = \mathrm{O}_n(\mathbb{R})$. Similarly, for an n-dimensional complex vector space with $\beta\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \beta(X,Y) = X^\top I_n \overline{Y}$ the standard sesquilinear form, the unitary group $\mathrm{U}_n(\mathbb{C})$ is naturally isomorphic to the group

$$\Phi^{-1}(\{g \in \mathrm{GL}(V)\colon (\forall v,w \in V)\ \beta(gv,gw) = \beta(v,w)\}.$$

One can also define indefinite unitary groups similar to what we did for orthogonal groups (see [HN12, p.23]).

1.3.3 Definition Consider the ***affine group*** $\mathrm{Aff}(V)$ of a vector space V defined as all the maps $V \to V$ of the form $\phi_{v,g}(x) = gx + v,\ g \in \mathrm{GL}(V), v \in V$. We write elements in $\mathrm{Aff}(V)$ simply as pairs $(v,g) \in V \times \mathrm{GL}_n(V)$. Then the composition in the group becomes

$$(v,g)(w,h) = (v+gw, gh),$$

$(0, \mathrm{id}_V)$ is the identity, and inversion is given by

$$(v,g)^{-1} = (-g^{-1}v, g^{-1})$$

For $V = \mathbb{K}^n$ we also write $\mathrm{Aff}_n(\mathbb{K}) = \mathrm{Aff}(\mathbb{K}^n)$, then

$$\Phi\colon \mathrm{Aff}_n(\mathbb{K}) \to \mathrm{GL}_{n+1}(\mathbb{K}),\ \Phi(v, g) = \begin{pmatrix} [g] & v \\ 0 & 1 \end{pmatrix}$$

is an injective group homomorphism, where $[g]$ denotes the matrix of the linear map with respect to the canonical standard basis of $\mathbb{K}^n$.

Let $V = \mathbb{R}^n$ and consider the euclidean metric $d(x, y) = \|x - y\|_2$ on $\mathbb{R}^n$. We define

$$\mathrm{Iso}_n(\mathbb{R}) = \{g \in \mathrm{Aff}_n(\mathbb{R})\colon (\forall x, y \in \mathbb{R}^n)\ d(gx, gy) = d(v, w)\}.$$

This is the group of *affine isometries* of the euclidian n-space.

Due to the Mazur-Ulam theorem[7], every bijective isometry of a normed space is an affine map, whence the requirement $g \in \mathrm{Aff}_n(\mathbb{R})$ in the definition of $\mathrm{Iso}_n(\mathbb{R})$ is superfluous.

1.3 Exercises

1. Let V be a finite dimensional $\mathbb{K}$-vector space and $\beta\colon V \times V \to \mathbb{K}$ a bilinear form.

a) Choose a basis for your favorite finite dimensional V and work out in this example the identification Φ and its inverse.

b) Show that β is (skew)symmetric if and only if its matrix B is (skew)symmetric (recall B is symmetric if $B^\top = B$ and skew-symmetric. if $B^\top = -B$).

c) Check which of the bilinear forms discussed in this section are non degenerate.

2. Let V be a finite dimensional complex vector space and $\beta\colon V \times V \to \mathbb{K}$ be a symmetric bilinear form.

a) Show that there exists an orthogonal basis $v_1, \dots, v_n$ with $\beta(v_j, v_j) = 1, j = 1, \dots, p$ and $\beta(v_j, v_j) = 0, j > p$.
Hint: Show that for a vector $v \in V$ with $\beta(v, v) = 0$ one has a decomposition $V = \mathrm{span}(v) \oplus \{v\}^\perp$. Here $\{v\}^\perp$ means the orthogonal complement, i.e. the subspace spanned by all $w \in V$ such that $\beta(v, w) = 0$. To prove this use Gram-Schmidt algorithm, then use induction.

b) Show that each invertible symmetric matrix $B \in \mathrm{GL}_n(\mathbb{C})$ can be written as $B = AA^\top$ for some $A \in \mathrm{GL}_n(\mathbb{C})$.

3. Let $\beta\colon V_1 \times V_2 \to W$ be a bilinear map between (finite dimensional) vector spaces V_1, V_2, W. Prove the product rule $\mathbf{d}\beta(v_1, v_2)(x_1, x_2) = \beta(v_1, x_2) + \beta(x_1, v_2)$.

[7] See e.g. https://arxiv.org/pdf/1306.2380 for a short proof.

1.4. The Matrix exponential

In this section we discuss the familiar matrix exponential and shall set it in the context of Lie theory. The main point will be to connect the matrix exponential to the 1-parameter groups which were discussed in the introduction. Further, we shall develope the Lie theoretic properties of the exponential. In later chapters these will help us to define the Lie algebra of a Lie group. Let us first say what we mean by a 1-parameter subgroup of a matrix Lie group.

1.4.1 Definition A 1-*parameter subgroup* in a linear Lie group $G \subseteq \mathrm{GL}_n(\mathbb{K})$ is a continuous group homomorphism $\phi\colon (\mathbb{R}, +) \to G \subseteq \mathrm{GL}_n(\mathbb{K})$.

Note that it is easier here to define the 1-parameter group as a morphism (and not as the image of this map) since it is in this way easier to explain how the group depends on the reals. Indeed if ϕ is a 1-parameter subgroup in $\mathrm{GL}_n(\mathbb{K})$, then

$$\phi(t+s) = \phi(t)\cdot\phi(s)(= \phi(s)\cdot\phi(t)), \qquad \forall s,t \in \mathbb{R}, \quad \phi(0) = I_n.$$

How do we find examples of such subgroups? First of all notice:

1.4.2 Example For $n = 1$ we have $\mathrm{GL}_n(\mathbb{R}) = \mathbb{R}^\times := \mathbb{R}\setminus\{0\}$, where matrix multiplication simply becomes multiplication of numbers. Then $\exp\colon \mathbb{R} \to \mathbb{R}^\times, \exp(t) = e^t$ is a 1-parameter group (which is even smooth).

Can we generate more examples this way? For this, we would like to insert a matrix $M \in M_n(\mathbb{K})$ into the exponential series

$$e^M := \sum_{k=0}^{\infty} \frac{M^k}{k!} \qquad \text{which we call the matrix exponential function.}$$

To make sense of this, we need to see that the series indeed converges (to some matrix). Set $S_\ell := S_\ell(M) := \sum_{k=0}^{\ell} M^k/k!$. Then we obtain a Cauchy sequence:

$$\|S_\ell - S_m\| \leq \sum_{k=m+1}^{\ell} \left\|\frac{M^k}{k!}\right\| \leq \sum_{k=m+1}^{\ell} \frac{\|M\|^k}{k!} < \varepsilon \text{ for } \ell, m \text{ large,}$$

as the right hand side is a partial sum of the exponential series (for the number $\|M\|$) and we know that this series converges. So $e^M = \lim_{k\to\infty} S_k(M)$ is again a matrix for every matrix M. Moreover, $\|e^M\| \leq e^{\|M\|}$ and the series giving the matrix exponential converges absolutely.

1.4.3 Lemma *For $X, Y \in M_n(\mathbb{K})$ such that $XY = YX$, we have $e^{X+Y} = e^X \cdot e^Y$.*

Proof. Since the matrices commute, the usual complex numbers proof works: We use the binomial theorem $(X+Y)^n = \sum_{k=0}^{n} \binom{n}{k} X^k Y^{n-k}$ and since the exponential series converges absolutely, we may resum as we please:

$$e^X e^Y = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{X^k}{k!} \frac{Y^{n-k}}{(n-k)!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} X^k Y^{n-k} = \sum_{n=0}^{\infty} \frac{(X+Y)^n}{n!} = e^{X+Y}.$$

□

If the matrices X, Y do not commute, there is no easy way to obtain a formula as in Lemma 1.4.3 for the product of exponentials. However, a closed power series formula called the Baker-Campbell-Hausdorff(-Dynkin) formula exists for $e^X e^Y$ (see Section 1.9 for more information).

1.4.4 Lemma *For $X \in M_n(\mathbb{K})$ and $C \in \mathrm{GL}_n(\mathbb{K})$ we have $e^{CXC^{-1}} = Ce^X C^{-1}$*

Proof. Since $(CXC^{-1})^n = CX^nC^{-1}$ the statement follows directly from the definition of e^X as a power series in X. □

Now lets apply what we learned from Lemma 1.4.3:

1.4.5 Lemma *For all $X \in M_n(\mathbb{K})$ and $t, s \in \mathbb{K}$ we have*

$$e^{tX} e^{sX} = e^{(t+s)X}, \quad (e^X)^{-1} = e^{-X}.$$

Proof. For every matrix $(tX)(sX) = (ts)X = (sX)(tX)$, for all $t, s \in \mathbb{K}$, whence the identity follows from Lemma 1.4.3. In particular $e^{-X} e^X = e^{X-X} = e^0 = I_n$. So every element e^X has inverse e^{-X}. □

The previous result shows that the matrix exponential actually generates a subgroup of $\mathrm{GL}_n(\mathbb{K})$ when we feed it the curve $t \to tX$. Let us see that the resulting map actually is differentiable, whence we get a 1-parameter curve which solves a certain differential equation.

1.4.6 Proposition *For every $X \in M_n(\mathbb{K})$, the function $\phi_X : \mathbb{R} \to \mathrm{GL}_n(\mathbb{K})$, $\phi_X(t) = e^{tX}$ is differentiable, and solves the initial value problem*

$$\frac{d}{dt}\phi_X(t) = Xe^{tX} = X\phi_X(t), \quad \phi_X(0) = I_n \tag{1.3}$$

In particular, $\left.\frac{d}{dt}\right|_{t=0} \phi_X(t) = X$ and $\phi_X(t)$ is a 1-parameter group in $\mathrm{GL}_n(\mathbb{K})$.

Proof. By definition we have $\phi_X(t) = \sum_{n=0}^{\infty} \frac{(tX)^n}{n!} = \sum_{n=0}^{\infty} \frac{X^n}{n!} t^n$. So every component of the resulting matrix is a power series in t. The radius of convergence of this power series is ∞ since it is dominated by the exponential series. Hence, ϕ_X is continuous in t (as all of its components are continuous in t) and we may differentiate the series in every component term by term to get the differential (Exercise: Work out the details!). That ϕ_X is a 1-parameter group thus follows from Lemma 1.4.5. We now obtain

$$\frac{d}{dt}\phi_X(t) = \frac{d}{dt}\sum_{n=0}^{\infty} \frac{(tX)^n}{n!} = \sum_{n=0}^{\infty} \frac{nt^{n-1}X^n}{n!} = X \sum_{n=0}^{\infty} \frac{(tX)^n}{n!} = X\phi_X(t)$$

Since $\phi_X(0) = e^0 = I_n$ the claim (1.3) follows. □

We actually gain a lot of information by the fact that the curves ϕ_X solve the initial value problem (1.3). For this we need to discuss briefly solution curves of vector fields (see Appendix A.3). For every $X \in M_n(\mathbb{K})$ we can define two special vector fields

$$X^r : M_n(\mathbb{K}) \to M_n(\mathbb{K}),\ X^r(M) = XM. \quad X^l : M_n(\mathbb{K}) \to M_n(\mathbb{K}),\ X^l(M) = MX. \tag{1.4}$$

Since matrix multiplication is smooth by the proof of Proposition 1.1.2, we immediately learn that both X^r and X^l are smooth vector fields for every X (the somewhat weird supscripts denote the direction by which we multiply X with other matrices, there is a deeper meaning to this, see Exercises). In vector field notation (1.3) becomes

$$\frac{d}{dt}\phi_X(t) = X^r(\phi_X(t)),$$

so ϕ_X is the integral curve starting at the identity of the vector field X^r. In particular, Proposition 1.4.6 shows that integral curves of the vector field X^r exist on all of $\mathbb{R}$. Thus we employ our knowledge of flows of vector fields to establish:

1.4.7 Proposition *The matrix exponential function*

$$\exp : M_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K}), \quad \exp(M) = e^M$$

is smooth with $\mathbf{d}\exp(0) = \mathrm{id}_{M_n(\mathbb{K})}$.

Proof. We have already seen that $\Psi : M_n(\mathbb{K}) \to C^\infty(M_n(\mathbb{K}), M_n(\mathbb{K})), \Psi(X) = X^r$ makes sense. The associated map $\Psi^\wedge : M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to M_n(\mathbb{K}), \Psi^\wedge(X, M) = XM = X^r(M)$ is clearly smooth by Proposition 1.1.2. Every one of the vector fields $\Psi(X) = X^r$ is complete by Exercise 1.4 4. d). Now Corollary A.3.14 yields a smooth map $\Phi : \mathbb{R} \times M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to M_n(\mathbb{K})$ such that $\Phi(t, X, I_n)$ is the integral curveof the vector field $\Psi(X) = X^r$ starting at I_n. By construction $\exp(X) = \Phi(1, X, I_n)$ is smooth.

For the derivative, we observe $\mathbf{d}\exp(0)(X) = \frac{d}{dt}\Big|_{t=0} \exp(tX) \overset{(1.3)}{=} X$. This establishes the formula. □

1.4.8 Remark For $n = 1$ the matrix exponential $\exp\colon M_1(\mathbb{R}) = \mathbb{R} \to \mathrm{GL}_1(\mathbb{R}) = \mathbb{R}^\times$, $\exp(x) = e^x$ is injective. This is not the case for $n > 1$. For example

$$\exp\left(\begin{pmatrix} 0 & -2\pi \\ 2\pi & 0 \end{pmatrix}\right) = I_2$$

since

$$\exp\left(\begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix}\right) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}, \quad \forall t \in \mathbb{R}.$$

This example is just the real picture of the relation $e^{2\pi i} = 1$.

1.4.9 Proposition *For each sufficiently small open neighborhood U of 0 in $M_n(\mathbb{K})$, the map*

$$\exp|_U : U \to \mathrm{GL}_n(\mathbb{K})$$

is a diffeomorphism onto an open neighborhood of I_n in $\mathrm{GL}_n(\mathbb{K})$.

Proof. As the derivative of exp in 0 is the identity (i.e. an invertible matrix) by Proposition 1.4.7, the inverse function Theorem A.2.7 shows that exp admits a smooth inverse function near 0. □

1.4.10 Definition If U is as in Proposition 1.4.9 and $V = \exp(U)$, we define $\log_V : V \to U \subseteq M_n(\mathbb{K})$, $M \mapsto \exp|_V^{-1}(M)$. We call log also the *matrix logarithm*.

One can show that the matrix logarithm can be computed by inserting matrices near the identity matrix into the Taylor series describing the logarithm. Thus it is justified to call the inverse of the matrix exponential the logarithm. See [HN12] for details. We will now use matrix exponential and logarithm to establish some results for linear Lie groups. The proofs of the following three results all start by choosing neighborhoods on which exp is a diffeomorphism.

1.4.11 Theorem (1-parameter group theorem) *Every 1-parameter group $\gamma\colon \mathbb{R} \to \mathrm{GL}_n(\mathbb{K})$ is of the form $\gamma(t) = e^{tX}$ for some $X \in M_n(\mathbb{K})$. In particular, every 1-parameter group is smooth and completely determined by its derivative in 0.*

Proof. Let $\gamma\colon \mathbb{R} \to \mathrm{GL}_n(\mathbb{K})$ be a 1-parameter group. That γ is uniquely determined by its derivative in 0 follows directly from Proposition 1.4.6 and the uniqueness of solutions to differential equations, Theorem A.3.5, if we can prove that $\gamma(t) = e^{tX}$.

To see this, pick U a convex[8] and symmetric (i.e. $U = -U$) 0-neighborhood in $M_n(\mathbb{K})$ such that exp restricts to a diffeomorphism on U (it exists by shrinking the neighborhood in Proposition 1.4.9). Define $U_1 := \frac{1}{2}U$. Since γ is continuous in 0 there exists $\varepsilon > 0$ such that $\gamma([-\varepsilon,\varepsilon]) \subseteq \exp(U_1)$. Then $\alpha\colon [-\varepsilon,\varepsilon] \to U_1, \alpha(t) := \log_U(\gamma(t))$ is a continuous curve with $\exp(\alpha(t)) = \gamma(t)$ for $|t| \leq \varepsilon$. For any such t we have

$$\exp\left(2\alpha\left(\frac{t}{2}\right)\right) = \exp\left(\alpha\left(\frac{t}{2}\right)\right)^2 = \gamma\left(\frac{t}{2}\right)^2 = \gamma(t) = \exp(\alpha(t)).$$

As exp is injective on U we see that $\alpha(t/2) = \alpha(t)/2$ for $|t| \leq \varepsilon$. Inductively we obtain

$$\alpha\left(\frac{t}{2^k}\right) = \frac{1}{2^k}\alpha(t), \quad \text{for } |t| \leq \varepsilon, k \in \mathbb{N}. \tag{1.5}$$

Further, $\alpha(t) \in \frac{1}{2^k}U_1$ for $|t| \leq \varepsilon/2^k$. For $m \in \mathbb{Z}$ with $|m| \leq 2^k$ and $|t| \leq \varepsilon/2^k$ we now have $m\alpha(t) \in U_1$ and

$$\exp(m\alpha(t)) = \gamma(t)^m = \gamma(mt) = \exp(\alpha(mt)).$$

Exploiting again that exp is injective on U_1, we obtain

$$\alpha(mt) = m\alpha(t) \text{ for } m \leq 2^k, |t| \leq \varepsilon/2^k \tag{1.6}$$

Combining (1.5) and (1.6) we obtain

$$\alpha\left(\frac{m}{2^k}t\right) = \frac{m}{2^k}\alpha(t) \text{ for } |t| \leq \varepsilon, k \in \mathbb{N}, m \in \mathbb{Z}, |m| \leq 2^k.$$

Since the numbers $\frac{mt}{2^k}, m \in \mathbb{Z}, k \in \mathbb{N}, |m| \leq 2^k$ are dense in the interval $[-t,t]$, the continuity of α implies that

$$\alpha(t) = \frac{t}{\varepsilon}\alpha(\varepsilon) \quad \text{for } |t| \leq \varepsilon.$$

In particular, α is smooth and of the form $\alpha(t) = tX$ for some $x \in M_n(\mathbb{K})$. Hence $\gamma(t) = \exp(tX)$ for $|t| \leq \varepsilon$, but then $\gamma(dt) = \exp(dtx)$ for $d \in \mathbb{N}$ leads to $\gamma(t) = \exp(tX)$ for each $t \in \mathbb{R}$. □

1.4.12 Theorem (No small subgroup Theorem) *There exists an open neighborhood V of I_n in $\mathrm{GL}_n(\mathbb{K})$ such that $\{I_n\}$ is the only subgroup of $\mathrm{GL}_n(\mathbb{K})$ contained in V.*

Proof. Pick an open neighborhood U as in Proposition 1.4.9. Shrinking U we may assume that U is bounded and convex. We set $U_1 := \frac{1}{2}U$. Let $G \subseteq V := \exp(U_1)$ be a subgroup of $\mathrm{GL}_n(\mathbb{K})$ and $g \in G$. Then we can write $g = \exp(x)$ with $x \in U_1$ and assume

[8] Recall a set C in a vector space is convex if for every pair $x, y \in C$ the convex combinations $(1-t)x + ty, t \in [0,1]$ are contained in C.

that $x = 0$. Let $k \in \mathbb{N}$ be maximal with $kx \in U_1$ (this is possible since U is bounded!). Then

$$g^{k+1} = \exp((k+1)x) \in G \cap V$$

implies the existence of $y \in U_1$ with $\exp((k+1)x) = \exp(y)$ Since $(k+1)x \in 2U_1 = U$ follows from $\frac{k+1}{2}x \in [0,k]x \subseteq U_1$, and $\exp|_U$ is injective, we obtain $(k+1)x = y \in U_1$ contradicting the maximality of k. Therefore $g = I_n$. □

1.4.13 Proposition ((Lie-)Trotter product formula) *For any $X, Y \in M_n(\mathbb{K})$ we have*

$$\exp(X+Y) = \lim_{k\to\infty} \left(\exp\left(\frac{X}{k}\right) \exp\left(\frac{Y}{k}\right) \right)^k .$$

Proof. Let again U be a neighborhood of 0 in $M_n(\mathbb{K})$ as in Proposition 1.4.9 for which $\exp|_U$ is a diffeomorphism onto $V = \exp(U)$. Define

$$U^{[2]} := \{(x,y) \in U \times U : \exp(x)\exp(y) \in V\} .$$

Then $U^{[2]} = (m \circ (\exp|_U \times \exp|_U))^{-1}(V) \cap U \times U$ is open as the preimage of the open set V under a continuous map and contains $(0,0)$. For $(x,y) \in U^{[2]}$ we can then define

$$x \star y := \log_V(\exp(x)\exp(y)).$$

and $\mu\colon U^{[2]} \to M_n(\mathbb{K}), \mu(x,y) := x \star y$ is smooth. Since $\mu(x,0) = x = \mu(0,x)$, we have

$$\mathbf{d}\mu(0,0)(x,y) = \mathbf{d}\mu(0,0)(x,0) + \mathbf{d}\mu(0,0)(0,y) = x + y.$$

Now pick $X, Y \in M_n(\mathbb{K})$ and choose $k \in \mathbb{N}$ so large that the line segment connecting $(0,0)$ with $(\frac{1}{k}X, \frac{1}{k}Y)$ is in $U^{[2]}$. Now the mean value theorem Proposition A.2.10 implies

$$\begin{aligned}
\lim_{k\to\infty} k \cdot \left(\frac{1}{k}X \star \frac{1}{k}Y\right) &= \lim_{k\to\infty} k \cdot \left(\mu\left(\frac{1}{k}X, \frac{1}{k}Y\right) - \underbrace{\mu(0,0)}_{=0}\right) \\
&= \lim_{k\to\infty} k \int_0^1 \mathbf{d}\mu\left((0,0) + t\left(\frac{1}{k}X, \frac{1}{k}Y\right)\right)\left(\frac{1}{k}X, \frac{1}{k}Y\right) \mathrm{d}t \\
&= \lim_{k\to\infty} \int_0^1 \mathbf{d}\mu\left(\frac{t}{k}(X,Y)\right)(X,Y)\,\mathrm{d}t = \mathbf{d}\mu(0,0)(X,Y) = X + Y
\end{aligned}$$

In the last line, we used Lemma A.2.9 to exchange the limit with the integral, as the integrand is continuous in the parameter k. Applying exp to both sides of the equation we get with $\exp(x \star y) = \exp(x)\exp(y)$ that:

$$\begin{aligned}
\exp(X+Y) &= \exp\left(\lim_{k\to\infty} k \cdot \left(\frac{1}{k}X \star \frac{1}{k}Y\right)\right) \\
&= \lim_{k\to\infty} \exp\left(\frac{1}{k}X \star \frac{1}{k}Y\right)^k = \lim_{k\to\infty} \left(\exp\left(\frac{1}{k}X\right)\exp\left(\frac{1}{k}Y\right)\right)^k
\end{aligned}$$

□

In the next section we will use exp to finally define the Lie algebra of a linear Lie group.

1.4 Exercises

1. Let M be a nilpotent matrix, i.e. there is $n \in \mathbb{N}$ with $M^n = 0$. How does this simplify the formula for e^M? Compute e^M explicitly for various 3×3-matrices which are strictly upper triangular (i.e. have only non-zero entries above the main diagonal).

2. Show that for all $X \in M_n(\mathbb{K})$ we have the identities

$$e^{X^T} = (e^X)^T, e^{\overline{X}} = \overline{e^X} \text{ and } e^{X^*} = (e^X)^* .$$

3. Let $M \in M_n(\mathbb{K})$ be a matrix which is diagonalisable with eigenvalues $\lambda_1, \dots, \lambda_n$. Show that e^M is diagonalisable with eigenvalues $e^{\lambda_1}, \dots, e^{\lambda_n}$.

4. Fix $M, N \in M_n(\mathbb{K})$ and denote by $\lambda_M, \rho_N : M_n(\mathbb{K}) \to M_n(\mathbb{K})$ the left (right) multiplication $\lambda_M(X) = MX$ and $\rho_N(X) = XN$.

a) Show that $\lambda_M \circ \rho_N = \rho_N \circ \lambda_M$. Why does this not imply that M and N commute?

b) Compute the maps[9] $\mathbf{d}\lambda_M : M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to M_n(\mathbb{K}), (X, H) \mapsto \mathbf{d}\lambda_M(X)(H)$ and $\mathbf{d}\rho_N : M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to M_n(\mathbb{K}), (X, H) \mapsto \mathbf{d}\rho_M(X)(H)$.

c) Let $X \in M_n(\mathbb{K})$ and $X^r(A) := XA$ and $X^\ell(A) := AX$ be the smooth vector fields from (1.4). Prove that $X^r \circ \rho_N(A) = \mathbf{d}\rho_N(A) \circ X^r(A)$ and $X^\ell \circ \lambda_M(A) = \mathbf{d}\lambda_M(A) \circ X^\ell(A)$.
This property is called *right/left invariance* and we have thus proved that X^ℓ is a left invariant vector field, while X^r is a right invariant vector field (this explains the weird r, ℓ supscripts).

d) By Proposition 1.4.6 the maximal integral curve ϕ_X of X^r starting at I_n is defined on $\mathbb{R}$. Show that $\phi_X^M(t) := \phi_X(t)M$ is the integral curve of X^r starting at M and deduce that X^r is a complete vector field.

5. Explain the trick by which we obtained continuity of e^{tX} in t and calculated the differential of e^{tX} in the proof of Proposition 1.4.6. For explanation of $\mathbf{d}\exp$ visit: https://gowers.wordpress.com/2014/02/22/differentiating-power-series/.

6. Prove that for $X, Y \in M_n(\mathbb{K})$ with $XY = YX$ the matrix exponential satisfies $\exp(X)Y = Y\exp(X)$. Show then that a 1-parameter group $\phi_X(t) = e^{tX}$ solves the differential equation $\frac{d}{dt}\phi_X(t) = \phi_X(t)X$.
Hint: Use Proposition 1.4.6 and commutativity. For the second part consider (1.3).

7. Convince yourself that the computation of the exponential in Remark 1.4.8 is correct.

[9] There is a slight abuse of notation as $\mathbf{d}f$ usually means the map taking values in the linear operators.

1.5. The Lie algebra of a linear Lie group

Recall that we have defined a linear Lie group as a closed subgroup of $\mathrm{GL}_n(\mathbb{K})$. In this section we define a new object attached to a linear Lie group, the Lie algebra.

1.5.1 Definition Let $G \subseteq \mathrm{GL}_n(\mathbb{K})$ be a linear Lie group. Then we define

$$\mathbf{L}(G) := \{X \in M_n(\mathbb{K}) \mid \forall t \in \mathbb{R}, \ \exp(tX) \in G\}$$

and call this set the *Lie algebra* of G.

For an arbitrary finite dimensional vector space $V = \mathbb{R}^n$, we have $\mathrm{GL}(V) = \mathrm{GL}_n(\mathbb{K})$ and hence can also assign a Lie algebra $\mathbf{L}(G)$ to closed subgroups of $\mathrm{GL}(V)$.

Note that by Proposition 1.4.6 for each matrix $X \in M_n(\mathbb{K})$, the map $\phi_X(t) = \exp(tX)$ is a 1-parameter group with $\frac{d}{dt}\Big|_{t=0} \phi_X(t) = X$. Hence the Lie algebra $\mathbf{L}(G)$ of a linear Lie group G collects all matrices which appear as derivatives of 1-parameter groups which are contained in G. In a sense we can thus think of the Lie algebra as derivatives at the identity (since every 1-parameter group takes 0 to the identity matrix). We will come back later to this observation.

1.5.2 Lemma *The Lie algebra $\mathbf{L}(G)$ of a linear Lie group $G \subseteq \mathrm{GL}_n(\mathbb{K})$ is a $\mathbb{R}$-vector space.*

Proof. As $\exp(t0) = I_n \in G$ for all $t \in \mathbb{R}$, $\mathbf{L}(G)$ contains at least 0. If $x \in \mathbf{L}(G)$, then $sx \in \mathbf{L}(G)$ since $\exp(tsx) = \exp((st)x) \in G$ by definition.

Let now $x, y \in \mathbf{L}(G)$. For $k \in \mathbb{N}$ and $t \in \mathbb{R}$ we have $\exp(\frac{t}{k}x), \exp(\frac{t}{k}y) \in G$ by definition. Now the Trotter product formula Proposition 1.4.13 yields

$$\exp(t(x+y)) = \lim_{k\to\infty} \left(\exp\left(\frac{t}{k}x\right)\exp\left(\frac{t}{k}y\right)\right)^k \in G,$$

because G is a subgroup and closed (whence the limit stays in G!). Therefore, $x + y \in \mathbf{L}(G)$. □

1.5.3 Remark In general $\mathbf{L}(G)$ will not be closed under multiplication with complex numbers, so it is not a $\mathbb{C}$-vector space in general and it will not be a Lie algebra over the complex numbers (see Definition 1.5.6 below). If $\mathbf{L}(G)$ has this additional property, the Lie group G is called a *complex Lie group*. One can show that then G actually is a complex manifold. In this lecture we will only consider real Lie algebras.

As a vector space $\mathbf{L}(G)$ does not inherit information on the group structure as it is not compatible with matrix multiplication, In general products Lie algebra elements will not be contained in the Lie algebra. However, the following takes the Lie algebra to itself:

1.5.4 Lemma *Let G be a linear Lie group and $x \in G$, then we define a linear map* $\mathrm{Ad}(x)\colon \mathbf{L}(G) \to \mathbf{L}(G)$, $\mathrm{Ad}(x)(v) = xvx^{-1}$. *The resulting map*

$$\mathrm{Ad}_G\colon G \to \mathrm{GL}(\mathbf{L}(G)), \quad g \to \mathrm{Ad}(g)$$

is called the adjoint representation *of the linear Lie group on its Lie algebra.*

Proof. We need to prove that $\mathrm{Ad}(x)$ defines a map from the Lie algebra to itself. For this let $v \in \mathbf{L}(G)$, then Lemma 1.4.4 yield for $t \in \mathbb{R}$ that

$$\exp(t(\mathrm{Ad}(x)(v)) = \exp(x(tv)x^{-1}) = x\exp(tv)x^{-1}$$

Now $\exp(tv) \in G$ by definition of the Lie algebra and since G is a subgroup and $x \in G$ the product on the right hand side is again in G. We deduce that $\mathrm{Ad}(x)\colon \mathbf{L}(G) \to \mathbf{L}(G)$ makes sense and defines a linear map on the Lie algebra. □

That the Lie algebra is stable under the adjoint representation is nice, but unfortunately this does not capture well the multiplicative structure as something which is defined on the Lie algebra only (we need to know G to compute Ad_G). Note that the adjoint representation captures non-commutativity of the Lie group in some sense. We will now define a structure which is intrinsic to the Lie algebra and does the same job.

1.5.5 Definition Let $A, B \in M_n(\mathbb{K})$ then we define the *commutator* as

$$[A, B] = AB - BA.$$

The commutator measures the failure of matrices to commute. We shall see soon that it preserves the Lie algebra of a linear Lie group. Let us define first abstract Lie algebras.

1.5.6 Definition (Abstract Lie algebra) Let L be a $\mathbb{K}$-vector space. A $\mathbb{K}$-bilinear map $[\cdot,\cdot]\colon L \times L \to L$ is called a *Lie bracket* (over $\mathbb{K}$) if

(L1) $[x, x] = 0$ for $x \in L$ and

(L2) $[x, [y, z]] = [[x, y], z] + [y, [x, z]]$ for $x, y, z \in L$ (Jacobi identity)

A *Lie algebra* (over $\mathbb{K}$) is a $\mathbb{K}$-vector space L, endowed with a Lie bracket. A subspace $E \subseteq L$ of a Lie algebra is called a *subalgebra* if $[E, E] \subseteq E$. A *homomorphism* $\phi\colon L_1 \to L_2$ of Lie algebras is a linear map with $\phi([x, y]) = [\phi(x), \phi(y)]$ for $x, y \in L_1$. A Lie algebra is said to be *abelian* of $[x, y] = 0$ holds for all $x, y \in L$.

1.5.7 Remark In many texts the Jacobi identity (L2) is stated in the alternative form

$$\sum_{\text{cycl}} [x, [y, z]] := [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \tag{1.7}$$

Here the "cycl" stands for "cyclic permutation of the x, y, z (see the arguments of the brackets!). The advantage is that the Jacobi identity is in this form quite easy to memorize. We leave it as an exercise to prove that (1.7) is equivalent to (L2).

Note first, that every vector space becomes trivially an abelian Lie algebra if we endow it with the bilinear map which is constant equal to 0. This is not very interesting, but the next result actually yields a source for interesting examples:

1.5.8 Example Let V be a $\mathbb{K}$-vector space and $\mathrm{Lin}(V, V)$ be the algebra of $\mathbb{K}$-linear maps from V to itself (where the multiplication is composition of linear maps). Define the commutator bracket

$$[F, G] = F \circ G - G \circ F$$

Then $(\mathrm{Lin}(V, V), [\cdot, \cdot])$ is a Lie algebra over $\mathbb{K}$ (Exercise!).

1.5.9 Proposition *For every linear Lie group G, the commutator bracket turns $\mathbf{L}(G)$ into a Lie algebra over $\mathbb{R}$.*

Proof. We leave it as an exercise to prove that the commutator actually is a Lie bracket. Instead we only show that it restricts to a bilinear map on the Lie algebra. For this we pick $v, w \in \mathbf{L}(G)$ and will now prove that then also $[v, w] \in \mathbf{L}(G)$. Consider the derivative of the adjoint action:

$$\frac{d}{dt}\mathrm{Ad}(e^{tv})(w) = \frac{d}{dt}e^{tv}we^{-tv} = ve^{tv}we^{-tv} + e^{tv}w(-v)e^{-tv}.$$

Here we used (1.1) to compute the derivative of the multiplication and the chain rule. Setting $t = 0$ the right-hand side becomes $vw - wv = [v, w]$. Thus

$$[v, w] = \left.\frac{d}{dt}\right|_{t=0} e^{tv}we^{-tv} = \lim_{t \to 0} \frac{1}{t}\left(e^{tv}we^{-tv} - w\right) = \lim_{t \to 0} \frac{1}{t}\left(\mathrm{Ad}(e^{tv})(w) - w\right). \tag{1.8}$$

Since $\mathbf{L}(G)$ is a vector space by Lemma 1.5.2 we deduce from Lemma 1.5.4 that for every $t \neq 0$ we have $\frac{1}{t}\left(\mathrm{Ad}(e^{tv})(w) - w\right) \in \mathbf{L}(G)$. As finite dimensional R-vector subspace, $\mathbf{L}(G)$ is a closed subset, whence the limit $[v, w]$ is contained in $\mathbf{L}(G)$. □

The Lie algebra of a linear Lie group does not depend on the special realisation of the linear Lie group as a group of matrices. This follows from the following Lemma (which is a reformulation of the 1-parameter group Theorem 1.4.11 for 1-parameter groups with values in G), the Trotter product formula and the so called commutator formular (see [HN12, Proposition 3.4.7]):

1.5.10 Lemma *Let $G \subseteq \mathrm{GL}_n(\mathbb{K})$ be a linear Lie group. If $\mathrm{Hom}(\mathbb{R}, G)$ denotes the set of all continuous group homomorphisms $(\mathbb{R}, +) \to G$, then the map*

$$\Gamma\colon \mathbf{L}(G) \to \mathrm{Hom}(\mathbb{R}, G), \quad \Gamma(x) = \gamma_x \text{ with } \gamma_x(t) := \exp(tx)$$

is a bijection

1.5 Exercises

1. Show that the commutator bracket $[A, B] = AB - BA$ is a Lie bracket and deduce that $(M_n(\mathbb{K}), [\cdot,\cdot])$ is a Lie algebra.

2. Let $(L, [\cdot,\cdot])$ be an abstract Lie algebra over the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Prove that

a) the Lie bracket is skew symmetric, i.e. $[x, y] = -[y, x]$ for all $x, y \in L$.

b) the Jacobi idenitity is equivalent to (1.7): $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.

3. Show that the commutator bracket $[F, G] = F \circ G - G \circ F$ for Example 1.5.8 induces a Lie bracket on $\mathrm{Lin}(V, V)$. How is this related to Exercise 1.5.1?

4. Explain how Lemma 1.5.10 follows from Theorem 1.4.11.

5. Let V, W be $\mathbb{K}$-vector spaces and $q\colon V \times V \to W$ a skew-symmetric bilinear map (i.e. q is bilinear with $q(v_1, v_2) = -q(v_2, v_1)$). Show that

$$[(v, w), (v', w')] := \left(0, q(v, v')\right)$$

is a Lie bracket on $\mathfrak{g} := V \times W$ with $x, y, z \in \mathfrak{g}$, we have $[x, [y, z]] = 0$.

6. Let $\mathfrak{g}$ be a two dimensional Lie algebra over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

a) Show that $\mathfrak{g}$ is either abelian or posesses a base $\{X, Y\}$ such that $[X, Y] = X$.

b) Show that the subspace $\mathfrak{s} = \mathrm{span}\left\{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right\}$ is a 2-dimensional Lie subalgebra of $M_2(\mathbb{R})$ (with the commutator bracket) which is not abelian. Construct a basis with the properties from a) for this Lie algebra then.

7.

Define $S := \left\{\begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} : x, y, z \in \mathbb{R}\right\}$ and $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$

a) Show that S is a subalgebra of $M_3(\mathbb{K})$ with the commutator bracket.

b) Show that $(\mathbb{R}^3, \times)$ is a Lie algebra which is not abelian.

c) Prove that $\phi\colon S \to \mathbb{R}^3, \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is a Lie algebra homomorphism (and actually an isomorphism of Lie algebras).

1.6. Computation of Lie algebras for linear Lie groups

1.6.1 Example It is clear from the definition that $\mathbf{L}(\mathrm{GL}_n(\mathbb{K})) = M_n(\mathbb{K})$ with the commutator bracket. Similarly $\mathbf{L}(\mathrm{GL}(V)) = \mathrm{Lin}(V, V)$.

Often in the literature Lie algebras associated to matrix Lie groups are denoted by using Fraktur[10] letters:

$$\mathfrak{gl}_n(\mathbb{K}) := M_n(\mathbb{K}) = \mathbf{L}(\mathrm{GL}_n(\mathbb{K})), \quad \mathfrak{gl}(V) = \mathrm{Lin}(V, V).$$

1.6.2 Example As in Exercise 1.1.2. there is a group homomorphism

$$\Phi\colon \mathbb{K}^n \to \mathrm{GL}_{n+1}(\mathbb{K}), \quad x \mapsto \begin{pmatrix} I_n & x \\ 0 & 1 \end{pmatrix}.$$

which identifies $\mathbb{K}^n$ with a linear Lie group $G = \Phi(\mathbb{K}^n) \subseteq \mathrm{GL}_{n+1}(\mathbb{K})$. It is easy to see that Φ is a diffeomorphism (i.e. smooth with smooth inverse).

Hence 1-parameter groups $\mathbb{R} \to \mathbb{K}^n$ correspond bijectively to 1-parameter groups with values in G. We exploit this to compute the Lie algebra of G: If $\gamma\colon \mathbb{R} \to \mathbb{K}^n$ is a 1-parameter group, then $\gamma(nt) = n\gamma(t)$ for all $n \in \mathbb{Z}$ and $t \in \mathbb{R}$. Thus for all $q \in \mathbb{Q}$ we obtain $\gamma(q) = q\gamma(1)$ and by continuity of γ this implies $\gamma(t) = t\gamma(1)$, $t \in \mathbb{R}$. Obviously we can freely choose $\gamma(1)$, so all smooth 1-parameter groups in G are of the form $\begin{pmatrix} I_n & tx \\ 0 & 1 \end{pmatrix}$. To find $\mathbf{L}(G)$ we use that by Lemma 1.5.10 it suffices to compute derivatives of 1-parameter groups in G (as these are of the form $\gamma_x(t) = e^{tx}$, where $x \in \mathbf{L}(G)$ is the derivative of the 1-parameter group at $t = 0$). We find that

$$\mathbf{L}(G) = \left\{ \left.\frac{d}{dt}\right|_{t=0} \begin{pmatrix} I_n & x \\ 0 & 1 \end{pmatrix} : x \in \mathbb{K}^n \right\} = \left\{ \begin{pmatrix} \mathbf{0} & x \\ 0 & 0 \end{pmatrix} : x \in \mathbb{K}^n \right\}.$$

Note that the commutator of two matrices in the Lie algebra is 0, whence $\mathbf{L}(G)$ is abelian (and note that it is no accident that also the group G is abelian!)

[10] While this looks nice, nobody bothers to ever write this on a blackboard, whence in particular the lecturer instead just underlines the small letters in case we want to express that we are talking about the Lie algebra... $\underline{gl}_n(\mathbb{K}) = \mathfrak{gl}_n(\mathbb{K})$).

The proof of the following Lemma is left as Exercise 1.6 1.:

1.6.3 Lemma *Let $X \in M_n(\mathbb{K})$ and denote by $\operatorname{tr} X = \sum_{i=1}^n X_{ii}$ its trace. Then*

$$\mathbf{d}\det(I_n)(X) = \operatorname{tr} X.$$

1.6.4 Example The group $G := \mathrm{SL}_n(\mathbb{K}) = \det^{-1}(1)$ is a linear Lie group. Consider $X \in M_n(\mathbb{K})$, then the curve $\gamma\colon \mathbb{R} \to \mathbb{K}^\times = \mathrm{GL}_1(\mathbb{K}), t \mapsto \det(e^{tX})$ is a continuous group homomorphism. Hence, Theorem 1.4.11 shows that it is of the form $\gamma(t) = e^{at}$ for $a = \gamma'(0)$. On the other hand the chain rule implies

$$a = \gamma'(0) = \mathbf{d}\det(I_n)(\mathbf{d}\exp(0)(x)) \overset{\text{Lemma 1.6.3}}{=} \operatorname{tr} X.$$

We deduce that $\det(e^X) = e^{\operatorname{tr} X}, \forall X \in M_n(\mathbb{K})$, whence

$$\begin{aligned}\mathfrak{sl}_n(\mathbb{K}) = \mathbf{L}(\mathrm{SL}_n(\mathbb{K})) &= \left\{X \in M_n(\mathbb{K})\colon (\forall t \in \mathbb{R})\ 1 = \det(e^{tX}) = e^{t \operatorname{tr} X}\right\}\\ &= \{X \in M_n(\mathbb{K})\colon \operatorname{tr} X = 0\}\end{aligned}$$

1.6.5 Example (Lie algebra of a finite groups) We have seen in Example 1.1.12 that every finite group is a linear Lie group. However, since finite groups are topologically discrete sets, their Lie algebra is the 0-dimensional vector space $\{0\}$. Thus the Lie algebra records in this case no information about the group structure.

In the rest of this section we compute the Lie algebras for various linear Lie groups we encountered in Section 1.3. For this we recall that for a finite dimensional $\mathbb{K}$-vector space V, we can choose a basis which gives an isomorphism $\mathrm{Lin}(V, V) = \mathrm{Lin}(\mathbb{K}^n, \mathbb{K}^n)$ and $\mathrm{GL}(V) = \mathrm{GL}_n(\mathbb{K})$. In particular, we can use this identification to make sense of e^{tF} for some $F \in \mathrm{Lin}(V, V)$ (which just corresponds under the identification to the matrix exponential of the matrix corresponding to F).

1.6.6 Lemma *Let V, W be finite dimensional vector spaces and $\beta\colon V \times V \to W$ a bilinear map. For $(x, y) \in \mathrm{Lin}(V, V) \times \mathrm{Lin}(W, W)$ the following are equivalent (we use the shorthand $y.v = y(v)$ and $y^n = y.y^{n-1}$ if y is a linear map, to save us some brackets):*

(a) $e^{ty}.\beta(v, v') = \beta(e^{tx}.v, e^{tx}.v')$ *for all $t \in \mathbb{R}$ and all $v, v' \in V$*

(b) $y.\beta(v, v') = \beta(x.v, v') + \beta(v, x.v')$ *for all $v, v' \in V$.*

Proof. (a) $\Rightarrow$ (b) : Take the derivative at $t = 0$ of the relation in (a) to obtain with the chain rule and the product rule for bilinear maps (Exercise 1.3 3)

$$y.\beta(v, v') = (e^{0y}y).\beta(v, v') = \beta((e^{0x}x).v, v') + \beta(v, (e^{0x}x).v') = \beta(x.v, v') + \beta(v, x.v').$$

$(b) \Rightarrow (a)$: If (b) holds, then we obtain inductively for $y^n.\beta(v, v')$

$$y^{n-1}.(y.\beta(v,v')) = y^{n-1}.(\beta(x.v,v') + \beta(v,x.v')) = \cdots = \sum_{k=0}^{n} \binom{n}{k} \beta(x^k.v, x^{n-k}.v')$$

For the exponential series, this yields with the identity

$$\begin{aligned} e^y.\beta(v,v') &= \sum_{n=0}^{\infty} \frac{1}{n!} y^n.\beta(v,v') = \sum_{n=0}^{\infty} \frac{1}{n!} \Big(\sum_{k=0}^{n} \binom{n}{k} \beta(x^k.v, x^{n-k}.v') \Big) \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \beta\Big(\frac{1}{k!} x^k.v, \frac{1}{(n-k)!} x^{n-k}.v' \Big) \overset{(*)}{=} \beta\Big(\sum_{k=0}^{\infty} \frac{1}{k!} x^k.v, \sum_{m=0}^{\infty} \frac{1}{m!} x^m.v' \Big) \\ &= \beta(e^x.v, e^x.v') \end{aligned}$$

The equality $(*)$ is the Cauchy-product formula for absolutely convergent series [HN12, Exercise 3.1.3]. Since (b) also holds for (tx, ty), $t \in \mathbb{R}$, this completes the proof. □

1.6.7 Proposition *Let V, W be finite-dimensional vector spaces and $\beta\colon V \times V \to W$ a bilinear map. The Lie algebra for the group*

$$\mathrm{Aut}(V,\beta) = \{g \in \mathrm{GL}(V) \colon (\forall v, v' \in V)\ \beta(g.v, g.v') = \beta(v,v')\}$$

is $\mathfrak{aut}(V,\beta) := \mathbf{L}(\mathrm{Aut}(V,\beta)) = \{x \in \mathfrak{gl}(V) \colon (\forall v, v' \in V)\ \beta(x.v,v') + \beta(v,x.v') = 0\}$

Proof. $X \in \mathfrak{gl}(V)$, $e^{tX} \in \mathrm{Aut}(V,\beta)$ if and only if $(X, 0)$ satisfies (a) in Lemma 1.6.6. □

1.6.8 Example Let $B \in M_n(\mathbb{K}), \beta(v,w) := v^\top Bw$ and

$$G := \{g \in \mathrm{GL}_n(\mathbb{K}) \colon g^\top Bg = B\} = \mathrm{Aut}(\mathbb{K}^n, \beta).$$

Then Proposition 1.6.7 implies that

$$\begin{aligned} \mathbf{L}(G) &= \{X \in \mathfrak{gl}_n(\mathbb{K}) \colon (\forall v, v' \in \mathbb{K}^n) \beta(x.v, v') + \beta(v, x.v') = 0\} \\ &= \{X \in M_n(\mathbb{K}) \colon X^\top B + Bx = 0\}. \end{aligned}$$

In particular, we obtain

$$\mathfrak{o}_n(\mathbb{K}) := \mathbf{L}(\mathrm{O}_n(\mathbb{K})) = \{X \in \mathfrak{gl}_n(\mathbb{K}) \colon X^\top = -X\} =: \mathrm{Skew}_n(\mathbb{K})$$

$$\mathfrak{o}_{p,q}(\mathbb{K}) := \mathbf{L}(\mathrm{O}_{p,q}(\mathbb{K})) = \{X \in \mathfrak{gl}_{p+q}(\mathbb{K}) \colon X^\top I_{p,q} + I_{p,q}X = 0\}, \text{ and}$$

$$\mathfrak{sp}_n(\mathbb{K}) := \mathbf{L}(\mathrm{Sp}_{2n}(\mathbb{K})) = \{X \in \mathfrak{gl}_{2n}(\mathbb{K}) \colon X^\top B + BX = 0\}, \text{ where } B = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}.$$

Applying Proposition 1.6.7 with $V = \mathbb{C}^n$ and $W = \mathbb{C}$, considered as real vector spaces, we obtain for the standard hermitian form $\beta\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \beta(z,w) := w^* I_n z$:

$$\begin{aligned} \mathfrak{u}_n(\mathbb{C}) := \mathbf{L}(\mathrm{U}_n(\mathbb{C})) &= \{X \in \mathfrak{gl}_n(\mathbb{C}) \colon (\forall z, w \in \mathbb{C}^n)\ w^* I_n X z + w^* X^* I_n z = 0\} \\ &= \{X \in \mathfrak{gl}_n(\mathbb{C}) \colon X^* = -X\}. \end{aligned}$$

1.6.9 Example (The Lie algebra of $\mathrm{SU}_2(\mathbb{C})$) By definition $\mathrm{SU}_2(\mathbb{C}) = \mathrm{SL}_2(\mathbb{C}) \cap \mathrm{U}_2(\mathbb{C})$. From Exercise 1.6 4. and Example 1.6.4 together with Example 1.6.8 we obtain

$$\mathfrak{su}_2(\mathbb{C}) := \mathbf{L}(\mathrm{SU}_2(\mathbb{C})) = \mathfrak{sl}_2(\mathbb{C}) \cap \mathfrak{u}_2(\mathbb{C}) = \{X \in \mathfrak{gl}_2(\mathbb{C}) \colon \mathrm{tr}(X) = 0 \text{ and } X^* = -X\}.$$

It is easy to show the Lie algebra is generated as a real vector space by the matrices

$$\mathrm{i}\sigma_1 := \mathrm{i}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathrm{i}\sigma_2 := \mathrm{i}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \mathrm{i}\sigma_1 := \mathrm{i}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The σ_j are known as *Pauli matrices* and occur in the Pauli equation in quantum mechanics. Due to the relevance of the Pauli matrices in physics we treated here only $\mathfrak{su}_2(\mathbb{C})$, a similar argument identifies also $\mathfrak{su}_n(\mathbb{C}), n \in \mathbb{N}$.

1.6 Exercises

1. Establish Lemma 1.6.3 and show that $\mathrm{tr}\, X = \mathbf{d}\det(I_n)(X), X \in M_n(\mathbb{K})$.
Hint: We already know that det is smooth. You could use the Laplace formula for det to show that $\det(I_n + tX) = 1 + t\,\mathrm{tr}\, X + t^2(\cdots)$.

2. Use Proposition 1.6.7 to compute the Lie algebra of the unitary group $\mathrm{U}_n(\mathbb{K})$. For this we consider $V = \mathbb{C}^n$ and $W = \mathbb{C}$ as real vector spaces!

3. The affine group $\mathrm{Aff}_n(\mathbb{R})$ is a linear Lie group, cf. Definition 1.3.3. Determine the Lie algebra $\mathfrak{aff}_n(\mathbb{R}) = \mathbf{L}(\mathrm{Aff}_n(\mathbb{R}))$.
Hint: Compute 1-parameter groups and use Lemma 1.5.10. Reduce to 1-parameter groups in: 1. the subgroup of all elements with $v = 0$, 2. the subgroup with $g = \mathrm{id}_{\mathbb{K}^n}$, check Example 1.6.2.

4. Let G_1, G_2 be linear Lie groups. Show that also $G_1 \cap G_2$ is a linear Lie group with $\mathbf{L}(G_1 \cap G_2) = \mathbf{L}(G_1) \cap \mathbf{L}(G_2)$.

5. Show that $\mathfrak{sl}_2(\mathbb{K})$ is a 3-dimensional $\mathbb{K}$-vector space by constructing a basis. Then compute the Lie brackets for the basis and find a 2-dimensional Lie subalgebra.

6. Let $\mathrm{i}\sigma_j$ be the Pauli matrices in $\mathfrak{su}_2(\mathbb{C})$ from Example 1.6.9. Show that...

a) the matrices $\mathrm{i}\sigma_j$ generate $\mathfrak{su}_2(\mathbb{C})$.

b) the Lie algebras $\mathfrak{su}_2(\mathbb{C})$ and $(\mathbb{R}^3, \times)$ from Exercise 1.5 7 are isomorphic.
Hint: Compute Lie brackets of the Pauli matrices and consider linear maps sending $\frac{1}{2}\mathrm{i}\sigma_j$ to some choice of the standard unit vectors e_k of $\mathbb{R}^3$.

c) $\mathfrak{su}_2(\mathbb{C})$ can not be isomorphic to $\mathfrak{sl}_2(\mathbb{R})$.
Hint: Use b) to argue that $\mathfrak{su}_2(\mathbb{C})$ can not have a 2-dimensional Lie subalgebra.

1.7. The Lie group exponential and the unit component

In this section we will consider the restriction of the matrix exponential to the Lie algebra of a linear Lie group. We already know that the restriction takes its image in the linear Lie group, but it inherits even more of the properties of the matrix exponential

1.7.1 Definition Let $G \subseteq \mathrm{GL}_n(\mathbb{K})$ be a linear Lie group then the continuous map

$$\exp_G \colon \mathbf{L}(G) \to G, \quad X \mapsto e^X.$$

is called the *Lie group exponential* of G.

We prove now some technical properties which will turn out to be very useful:

1.7.2 Lemma *Let E, F be linear subspaces with $E \oplus F = M_n(\mathbb{K})$, then the map*

$$\Phi \colon E \times F \to \mathrm{GL}_n(\mathbb{K}), \quad \Phi(x,y) = \exp(x)\exp(y)$$

restrict to a diffeomorphism of a neighborhood of $(0,0)$ to an open I_n-neighborhood.

Proof. As the derivative is linear, $\mathbf{d}\Phi(0,0)(x,y) = \mathbf{d}\Phi(0,0)(x,0) + \mathbf{d}\Phi(0,0)(0,y)$. Let us compute the pieces of the derivative with Proposition 1.4.7:

$$\mathbf{d}\Phi(0,0)(x,0) = \left.\frac{d}{dt}\right|_{t=0} \exp(tx)\exp(0) = x.$$

A similar argument for the other derivative shows $\mathbf{d}\Phi(0,0)(x,y) = x + y$. Now $E \oplus F = M_n(\mathbb{K})$ means that the map $E \times F \to M_n(\mathbb{K}), (x,y) \mapsto x + y$ is a linear isomorphism. We have thus proved that $\mathbf{d}\Phi(0,0)$ is an invertible linear map, whence the claim follows from the Inverse Function Theorem A.2.7. □

1.7.3 Proposition *Let G be a linear Lie group and $\exp_G$ its Lie group exponential. There exists an open 0-neighborhood $V_1 \overset{\circ}{\subseteq} M_n(\mathbb{K})$ and an open I_n-neighborhood V_2 such that $\exp_G$ restricts to a homeomorphism $V_1 \cap \mathbf{L}(G) \to V_2 \cap G$.*

Proof. Pick $U \overset{\circ}{\subseteq} M_n(\mathbb{K})$ satisfying the requirements of Proposition 1.4.9. Then $W := \exp(U)$ is an open set and we have a smooth inverse $\log \colon W \to U$ of $\exp$. We have $\exp(U \cap \mathbf{L}(G)) \subseteq G$, but we need to prove that also $\log(W \cap G) \subseteq \mathbf{L}(G)$. Since the Lie algebra $\mathbf{L}(G)$ is a real subspace of $M_n(\mathbb{K}) = \mathbb{K}^{n^2}$ by Lemma 1.5.2. If $\mathbb{K} = \mathbb{C}$ identify $\mathbb{K}^{n^2} = \mathbb{R}^{(2n)^2}$. Pick a real orthogonal complement E of $\mathbf{L}(G)$ in $M_n(\mathbb{K})$. Then we write $Z \in M_n(\mathbb{K}) = E \oplus \mathbf{L}(G)$ as $Z = Z_{\mathbf{L}(G)} + Z^{\perp}$. Now apply Lemma 1.7.2 with E and $F := \mathbf{L}(G)$, then $\Phi(Z) = \exp(Z_{\mathbf{L}(G)})\exp(Z^{\perp})$ is smooth and invertible near 0.

Argue by contradiction and assume that for every open I_n-neighborhood $V \subset W$ there exists $X \in V \cap G$ such that $\log X \notin \mathbf{L}(G)$. We can thus pick a sequence $X_k \in G$ such that $\lim_{k\to\infty} X_k = I_n$ but $\log(X_k) \notin \mathbf{L}(G)$.

For k sufficiently large, X_k is in an open unit neighborhood on which an inverse of Φ exists. Hence, $X_k = \Phi(Z_k) = e^{Z^k_{\mathbf{L}(G)}} e^{Z_k}$ for unique $Z^k_{\mathbf{L}(G)} \in \mathbf{L}(G)$ and $Z_k \in E$. Both sequences of Z's converge to 0 as $k \to \infty$ and we must have $\|Z_k\| \neq 0$ for all k. Since $X_k, e^{Z^k_{\mathbf{L}(G)}} \in G$, we have that

$$e^{Z_k} = e^{-Z^k_{\mathbf{L}(G)}} X_k \in G. \tag{1.9}$$

Since $\|Z_k\| \neq 0$ we can consider the sequence $Z_k / \|Z_k\|$ which has a convergent subsequence as it is contained in the unit sphere of E (which is compact). Passing to the subsequence we obtain $Y = \lim_{k\to\infty} Z_k / \|Z_k\| \in E$ and $\|Y\| = 1$. We claim now that $Y \in \mathbf{L}(G)$, i.e. that $e^{tY} \in G$ for all $t \in \mathbb{R}$ which, if it is true, is a contradiction since $Y \neq 0$. As $e^{Z_k} \to I_m$ and the Z_k are contained in U we must have $\|Z_k\| \to 0$. For any $t \in \mathbb{R}$ there is thus a sequence of integers m_k such that $m_k \|Z_k\| \to t$. Then we have

$$\exp(Z_k)^{m_k} = \exp(m_k Z_k) = \exp\left(m_k \|Z_k\| \frac{Z_k}{\|Z_k\|} \right) \to \exp(tY), \quad \text{as } k \to \infty$$

However, (1.9) tells us that the left hand side is contained in G and since G is closed also the limit is contained in G. This gives the contradiction and we deduce that there must be an open I_n-neighborhood V_2 such that $\log(V_2 \cap G) \subset \mathbf{L}(G)$. Since log is continuous and an inverse to exp we can set $V_1 = \log(V_2)$ to finish the proof. □

A direct consequence of Proposition 1.7.3 will be that every linear Lie group is a submanifold of $\mathrm{GL}_n(\mathbb{K})$, but more on that later. We derive some other useful information from Proposition 1.7.3. Note first that every element in the image of the exponential function is connected to the unit by a continuous path. If $X = \exp_G(Y)$ with $Y \in \mathbf{L}(G)$, then $t \mapsto \exp_G(tY)$ is such a path. We study now the subset of all such elements.

1.7.4 Lemma *Let G be a topological group with unit $\mathbf{1}$, and define*

$$G_0 := \{g \in G \mid \exists \gamma\colon [0,1] \to G \text{ continuous}, c(0) = \mathbf{1}, c(1) = g\}$$

the path component of the unit *or* unit component. *Then G_0 is a normal subgroup*[11].

Proof. To see that G_0 is a subgroup, pick $g, h \in G_0$ and (continuous) paths $\gamma_g, \gamma_h\colon [0,1] \to G$ connecting the respective elements with the unit. Since the group multiplication and inversion is continuous, so are the paths $t \mapsto \gamma_g(t)\gamma_h(t)$ and $t \mapsto (\gamma_g(t))^{-1}$ connecting

[11] Recall that a subgroup H of a group G is *normal* if for every $g \in G$ we have equality of sets $gH = Hg$.

the identity to gh and to g^{-1}. So products and inverses of elements in G_0 are again contained in G_0, so it is a (topological) subgroup of G.

To see that G_0 is normal it suffices to show that for every $x \in G$ the conjugate xgx^{-1} is again contained in G_0 if g is contained in G_0. However, $t \to x\gamma_g(t)x^{-1}$ is a path such that $x\gamma_g(0)x^{-1} = x\mathbf{1}x^{-1} = \mathbf{1}$ and $x\gamma_g(1)x^{-1} = xgx^{-1}$. Thus $xgx^{-1} \in G_0$. □

1.7.5 Example (Unit components of linear Lie groups)

- The following groups are connected whence $\mathrm{GL}_n(\mathbb{C})_0 = \mathrm{GL}_n(\mathbb{C})$, $\mathrm{U}_n(\mathbb{C})_0 = \mathrm{U}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbb{C})_0 = \mathrm{SL}_n(\mathbb{C})$,
- $\mathrm{O}_n(\mathbb{R})_0 = \mathrm{SO}_n(\mathbb{R})$. In particular, $\mathrm{SO}_n(\mathbb{R})$ is a normal subgroup of $\mathrm{O}_n(\mathbb{R})$.

Indeed we can say more about the unit component of a linear Lie group. It is useful however, to state first a more general statement about subgroups of linear Lie groups.

1.7.6 Proposition (Open subgroups of linear Lie groups) *Let G be a linear Lie group and H be a subgroup of G. Then the following is equivalent:*

(a) *H contains a neighborhood of the unit.*

(b) *H is an open subset of G.*

(c) *H is open and closed in G.*

(d) *H contains the unit component G_0.*

In particular, if H satisfies one of the above conditions, it is a linear Lie group.

Proof. (a) $\Rightarrow$ (b): Since H contains a neighborhood of the unit I_n, we may choose an open set $I_n \in U \subseteq H$. To see that H is open it suffices to prove that H is a neighborhood for every $h \in H$. However, as G is a topological group, $hU := \{g \in G \mid g = hu, u \in U\}$ is open (Exercise B.1 2.) which is contained in the subgroup H and since $I_n \in U$ we see that $h \in hU \subseteq H \subseteq G$. Thus H is open.
(b) $\Rightarrow$ (c): We already know that H is open, to see that it is closed we prove that $G \setminus H$ is open. Let $a \in G \setminus H$. Then aH is an a-neighborhood in G as H is open. If $aH \cap H = \emptyset$ we find $ah_1 = h_2$ and equivalently $a = h_2h_1^{-1} \in H$ as H is a subgroup. This contradicts our choice of a, whence $aH \cap H = \emptyset$. Thus $G \setminus H$ is open, whence H closed.
(c) $\Rightarrow$ (d): Since H is open and closed, $H \cap G_0$ is both open and closed in G_0 and contains I_n, i.e. it is not empty. But G_0 is path-connected by construction, whence in particular connected (see Exercise A.1.4), so we must have $H \cap G_0 = G_0$. In other words, $G_0 \subseteq H$.
(d) $\Rightarrow$ (a): From Proposition 1.7.3 we know that there is an open subset $W \subseteq G_0$ which contains the unit. Hence $W \subseteq G_0 \subseteq H$ contains a unit neighborhood.

If H satisfies 3. it is a closed subgroup of $G \subseteq \mathrm{GL}_n(\mathbb{K})$. As G is a linear Lie group, so closed, H is a closed subgroup of $\mathrm{GL}_n(\mathbb{K})$, i.e. a linear Lie group. □

1.7.7 Corollary *The unit component G_0 of a linear Lie group G is an open subgroup and a linear Lie group. Further, $\mathbf{L}(G_0) = \mathbf{L}(G)$.*

Proof. We know that G_0 is a subgroup. It contains a unit neighborhood by Proposition 1.7.3, whence Proposition 1.7.6 implies that it is an open subgroup and a linear Lie group. We leave the statement concerning the Lie algebra as an exercise. □

1.7.8 Corollary *Let G be a linear Lie group. Every element $A \in G_0$ can be written as*

$$A = e^{X_1} e^{X_2} \cdots e^{X_k}, X_1, \dots, X_k \in \mathbf{L}(G), \text{ for some } k \in \mathbb{N}.$$

Proof. By Proposition 1.7.3 there is an open identity neighborhood $W \subseteq G_0$ such that every element in W can be written as e^X for some $X \in \mathbf{L}(G)$. Define $U := W \cap W^{-1} \subseteq G_0$, then $g \in U$ if and only if $g^{-1} \in U$. Further, U is open and every element in U is of the form e^X for some $X \in \mathbf{L}(G)$. Then $H := \bigcup_{n \in \mathbb{N}} U^n$ is an open subgroup of G, so by Proposition 1.7.6 we have $G_0 \subseteq H$. This proves the claim. □

1.7 Exercises

1. Let G be a linear Lie group with unit component G_0. Prove that

a) $\mathbf{L}(G_0) = \mathbf{L}(G)$. Use this to compute the Lie algebra of $\mathrm{SO}_n(\mathbb{K})$.

b) Prove that the canonical quotient map $q\colon G \to G/G_0$, $g \mapsto gG_0$ becomes a continuous group morphism if we endow G/G_0 with the discrete topology (i.e. the topology for which every singleton is an open set). The topological group G/G_0 is also known as the *group of connected components* of G.

2. Let $I_n \in O \subseteq G_0 \subseteq G$ be an open subset in a linear Lie group G. Prove that the group generated by O, i.e. the set containing all finite products and their inverses of elements in O, coincides with the unit component G_0.

3. Let G be a linear Lie group such that $\mathbf{L}(G)$ is abelian.

a) Show that G_0 is a commutative group.
Hint: Use Lemma 1.4.3 and Corollary 1.7.8.

b) Take a non-commutative finite group H with the discrete topology, denote by $\mathbf{1}_H$ its unit and consider the product group $\hat{G} := G \times H$. Show that $\hat{G}$ is a non-commutative linear Lie group with $\hat{G}_0 = G_0 \times \{\mathbf{1}_H\}$. Deduce that $\mathbf{L}(\hat{G}) = \mathbf{L}(G) \times \{0\}$ is abelian while $\hat{G}$ is not commutative, but its unit component is commutative by construction!.

1.8. Interplay of linear Lie groups and Lie algebras

In this section we take investigate the relation between linear Lie groups and their Lie algebra. First of all we establish how Lie group morphisms induce Lie algebra morphisms.

1.8.1 Proposition *Let G_1, G_2 be linear Lie groups and $\phi\colon G_1 \to G_2$ a continuous group homomorphism. Then the derivative*

$$\mathbf{L}(\phi)(x) := \frac{d}{dt}\Big|_{t=0} \phi(\exp_{G_1}(tx))$$

exists for each $x \in \mathbf{L}(G_1)$ and defines a morphism of Lie algebras $\mathbf{L}(\phi)\colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ which satisfies

$$\exp_{G_2} \circ \mathbf{L}(\phi) = \phi \circ \exp_{G_1} \tag{1.10}$$

Furthermore, $\mathbf{L}(\phi)$ is the uniquely determined linear map satisfying (1.10).

Proof. For $x \in \mathbf{L}(G_1)$ we define a 1-parameter subgroup $\phi_x(t) = \phi(e^{tx})$ in G_2. Then the 1-parameter group Theorem 1.4.11 implies that $\phi_x(t) = \exp_{G_2}(ty)$ for some $y \in \mathbf{L}(G_2)$. Applying Proposition 1.4.6 we see that $y = \frac{d}{dt}\big|_{t=0} \phi_x(t) = \mathbf{L}(\phi)(x)$. For $t = 1$ we obtain in particular (1.10):

$$\exp_{G_2}(\mathbf{L}(\phi)x) = \phi(\exp_{G_1}(x)).$$

$\mathbf{L}(\phi)$ is $\mathbb{R}$-linear: For $x \in \mathbf{L}(G_1)$, $t \in \mathbb{R}$ we get

$$\exp_{G_2}(s\mathbf{L}(\phi)(tx)) = \phi(\exp_{G_1}(stx)) = \exp_{G_2}(ts\mathbf{L}(\phi)(x)), \quad s \in \mathbb{R}$$

which leads to $\mathbf{L}(\phi)(tx) = t\mathbf{L}(\phi)(x)$. Since ϕ is continuous, the Trotter product formula Proposition 1.4.13 implies

$$\begin{aligned}
&\exp_{G_2}(\mathbf{L}(\phi)(x+y)) = \phi(\exp_{G_1}(x+y)) \\
&= \lim_{k\to\infty} \phi\left(\left(\exp_{G_1}\left(\frac{1}{k}x\right)\exp_{G_1}\left(\frac{1}{k}x\right)\right)^k\right) = \lim_{k\to\infty}\left(\phi(\exp_{G_1}(\frac{1}{k}x))\phi\left(\exp_{G_1}\left(\frac{1}{k}x\right)\right)\right)^k \\
&= \lim_{k\to\infty}\left(\exp_{G_2}\left(\frac{1}{k}\mathbf{L}(\phi)(x)\right)\exp_{G_1}\left(\frac{1}{k}\mathbf{L}(\phi)(x)\right)\right)^k = \exp_{G_2}\left(\mathbf{L}(\phi)(x) + \mathbf{L}(\phi)(y)\right)
\end{aligned}$$

for all $x, y \in \mathbf{L}(G_1)$. Therefore $\mathbf{L}(\phi)(x+y) = \mathbf{L}(\phi)(x) + \mathbf{L}(\phi)(y)$ because the same formula holds with tx and ty instead of x, y. Hence $\mathbf{L}(\phi)$ is $\mathbb{R}$-linear.

$\mathbf{L}(\phi)$ is a Lie algebra morphism We will use the formula

$$\mathbf{L}(\phi)(\mathrm{Ad}(g)(x)) = \mathrm{Ad}(\phi(g))(\mathbf{L}(\phi)(x)) \text{ for } g \in G_1 \text{ and } x \in \mathbf{L}(G_1). \tag{1.11}$$

The proof of (1.11) is left as Exercise 1.8 3. Now recall from (1.8) the formula $[x, y] = \frac{d}{dt}\Big|_{t=0} e^{tx}ye^{-tx}$. We know that $\mathbf{L}(\phi)$ is linear, whence it commutes with the derivative and we obtain

$$\begin{aligned}\mathbf{L}(\phi)([x, y]) &= \frac{d}{dt}\Big|_{t=0} \mathbf{L}(\phi)(e^{tx}ye^{-tx}) \overset{(1.11)}{=} \frac{d}{dt}\Big|_{t=0} \phi(e^{tx})\mathbf{L}(\phi)(y)\phi(e^{-tx})) \\ &= \frac{d}{dt}\Big|_{t=0} e^{t\mathbf{L}(\phi)(x)}\mathbf{L}(\phi)(y)e^{-t\mathbf{L}(\phi(x)}) = [\mathbf{L}(\phi)(x), \mathbf{L}(\phi)(y)]\end{aligned}$$

Uniqueness of the linear map: Assume now that $\psi \colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ is linear with $\exp_{G_2} \circ \psi = \phi \circ \exp_{G_1}$ satisfies

$$\phi \circ \exp_{G_1}(tx) = \exp_{G_2}(\psi(tx)) = \exp_{G_2}(t\psi(x))$$

and therefore

$$\mathbf{L}(\phi)x = \frac{d}{dt}\Big|_{t=0} \exp_{G_2}(t\psi(x)) = \psi(x).$$

This establishes uniqueness and finishes the proof. □

1.8.2 Remark The property (1.10) means that the following diagram commutes

$$\begin{array}{ccc} \mathbf{L}(G_1) & \xrightarrow{\mathbf{L}(\phi)} & \mathbf{L}(G_2) \\ \big\downarrow{\scriptstyle \exp_{G_1}} & & \big\downarrow{\scriptstyle \exp_{G_2}} \\ G_1 & \xrightarrow{\quad \phi \quad} & G_2 \end{array}$$

This property is also called "naturality of the Lie group exponential" as it means that the Lie group exponential transforms Lie algebra morphisms to Lie group morphisms.

1.8.3 Lemma *If G_1, G_2, G_3 are linear Lie groups and $\phi_1 \colon G_1 \to G_2$ and $\phi_2 \colon G_2 \to G_3$ are continous group homomorphisms, then*

$$\mathbf{L}(\mathrm{id}_G) = \mathrm{id}_{\mathbf{L}(G)}, \qquad \mathbf{L}(\phi_2 \circ \phi_1) = \mathbf{L}(\phi_2) \circ \mathbf{L}(\phi_1).$$

This is the functor property of $\mathbf{L}$.

We leave the proof of Lemma 1.8.3 as Exercise 1.8 2. In the language of category theory Lemma 1.8.3 means that $\mathbf{L}$ defines a functor from the category of linear Lie groups (with continuous group homomorphisms as morphisms) to the category of real Lie algebras. The functor $\mathbf{L}$ is also known as the *Lie functor*.

1.8.4 Corollary *If $\phi\colon G_1 \to G_2$ is an isomorphism of linear Lie groups, then $\mathbf{L}(\phi)$ is an isomorphism of Lie algebras.*

Proof. Since ϕ is an isomorphism, both ϕ and ϕ^{-1} are continuous group homomorphisms. The functor property then yields

$$\mathrm{id}_{\mathbf{L}(G_1)} = \mathbf{L}(\mathrm{id}_{G_1}) = \mathbf{L}(\phi^{-1} \circ \phi) = \mathbf{L}(\phi^{-1}) \circ \mathbf{L}(\phi)$$

and $\mathbf{L}(\phi) \circ \mathbf{L}(\phi^{-1}) = \mathrm{id}_{\mathbf{L}(G_2)}$. Hence $\mathbf{L}(\phi)$ is an isomorphism with $\mathbf{L}(\phi)^{-1} = \mathbf{L}(\phi^{-1})$. □

1.8.5 Definition If V is a vector space and G a group, then a homomorphism $\phi\colon G \to \mathrm{GL}(V)$ is called a *representation of G on V*.

If $(L, [\cdot,\cdot])$ is a Lie algebra, then a homomorphism of Lie algebras $\psi\colon L \to L(V, V)$ (with the Lie algebra structure from Example 1.5.8) is called *representation of L on V*.

1.8.6 Corollary *If $\phi\colon G \to \mathrm{GL}(V)$ is a continuous representation of the linear Lie group G, then $\mathbf{L}(\phi)\colon \mathbf{L}(G) \to \mathbf{L}(\mathrm{GL}(V)) = \mathrm{Lin}(V, V)$ is a representation of $\mathbf{L}(G)$ on V.*

The representation $\mathbf{L}(\phi)$ in Corollary 1.8.6 is called the *derived representation* This is motivated by the fact that for each $X \in \mathbf{L}(G)$ we have

$$\mathbf{L}(\phi)X = \frac{d}{dt}\Big|_{t=0} e^{t\mathbf{L}(\phi)X} = \frac{d}{dt}\Big|_{t=0} \phi(\exp(tX)). \tag{1.12}$$

1.8.7 Example Let G be a linear Lie group. Then G has the following representations

(a) The *trivial representation* on a vector space V given by $G \to \mathrm{GL}(V), g \to \mathrm{id}_V$. The derived representation of this representation is the zero map $\mathbf{L}(G) \to \mathfrak{gl}(V)$.

(b) As G is a subgroup of $\mathrm{GL}_n(\mathbb{K})$. The inclusion $G \to \mathrm{GL}_n(\mathbb{K})$ is a representation of G, sometimes also called the *standard representation* of G. The associated derived representation is simply the inclusion of the Lie algebras $\mathbf{L}(G) \to \mathfrak{gl}_n(\mathbb{K})$

(c) The adjoint representation $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathbf{L}(G))$ on its Lie algebra, see Lemma 1.5.4. We will see in Exercise 1.8 3. that the associated derived representation $\mathrm{ad} = \mathbf{L}(\mathrm{Ad})\colon \mathbf{L}(G) \to \mathrm{Lin}(\mathbf{L}(G), \mathbf{L}(G))$ is given by the Lie bracket of G

We will return to representations of Lie groups after establishing the abstract setting of Lie theory.

1.8.8 Proposition *Let G be a linear Lie group. Then $\mathbf{L}(G)$ is abelian, i.e. $[X, Y] = 0$, $\forall X, Y \in \mathbf{L}(G)$, if and only if G_0 is commutative, i.e. G_0 is an abelian group.*

Proof. If $\mathbf{L}(G)$ is abelian, Exercise 1.7 3. a) shows that G_0 is commutative. Assume conversely that G_0 is a commutative group. Then for every $g \in G_0$ the conjugation $c_g \colon G_0 \to G_0, c_g(h) = ghg^{-1}$ is the identity id_{G_0}. We now use Exercise 1.8 3. together with Corollary 1.8.4 which shows that $\mathrm{Ad}(g) = \mathbf{L}(c_g) = \mathbf{L}(\mathrm{id}_{G_0}) = \mathrm{id}_{\mathbf{L}(G)}$ (where we exploit that $\mathbf{L}(G) = \mathbf{L}(G_0)$!). Since the adjoint representation of G_0 is constant in g, $\mathbf{L}(\mathrm{Ad}) = 0$, whence $\mathrm{ad} = \mathbf{L}(\mathrm{Ad}) = 0$ by Exercise 1.8 3., i.e. $[X, Y] = \mathrm{ad}(X)(Y) = 0$ for all $X, Y \in \mathbf{L}(G_0) = \mathbf{L}(G)$, whence $\mathbf{L}(G)$ is abelian. □

The Lie algebra controls the unit component via the exponential. The next example shows that the Lie algebra and unit component might fail to capture essential structure of the group.

1.8.9 Example We return to the group $\mathrm{SU}_2(\mathbb{C})$. In $\mathrm{SU}_2(\mathbb{C})$, the circle group

$$T := \left\{ \begin{pmatrix} x & 0 \\ 0 & \overline{x} \end{pmatrix} : x \in \mathbb{C}, |x| = 1 \right\},$$

is an abelian, compact subgroup, whence it is a linear Lie group. Its normaliser

$$N(T) = \{X \in \mathrm{SU}_2(\mathbb{C}) \colon XTX^{-1} = T\}$$

is also a compact subgroup with two connected components T and the circle (not a group!)

$$T' := \left\{ \begin{pmatrix} 0 & y \\ -\overline{y} & 0 \end{pmatrix} : y \in \mathbb{C}, |x| = 1 \right\} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} T.$$

As $N(T)_0 = T$ and the unit component is a normal subgroup, $N(T)/T$ is a group which can be identified with the multiplicative group $\{\pm 1\}$ of two elements, Computations with the matrices show that $N(T)$ is highly non-commutative, yet this non-commutativity can neither be detected by the abelian group T nor the quotient $N(T)/T$ or the Lie algebra!

This problem is related to an interesting relation between $\mathrm{SU}_2(\mathbb{C})$ and rotations. With some work one can show that the morphisms given by $\mathrm{Ad}(g)$ for $g \in \mathrm{SU}_2(\mathbb{C})$ are rotations on $\mathfrak{su}_2(\mathbb{C})$. As we have seen in Example 1.6.9 this Lie algebra is a 3-dimensional real vector space, whence the adjoint representation yields a continuous group morphism $\mathrm{SU}_2(\mathbb{C}) \to \mathrm{SO}_3(\mathbb{R})$. This morphism yields a 2-fold covering of $\mathrm{SO}_3(\mathbb{R})$ as its kernel is $\{\pm I_2\}$. We do not discuss coverings of Lie groups in these notes and refer to [DK00, 1.2.B] for more information.

1.8 Exercises

1. Show that for a continous group homomorphism $\phi \colon G_1 \to G_2$ between linear Lie groups, the map $\mathbf{L}(\phi)$ satisfies for the adjoint representation from Lemma 1.5.4 the identity

$$\mathbf{L}(\phi)(\mathrm{Ad}(g)(x)) = \mathrm{Ad}(\phi(g))(\mathbf{L}(\phi)(x)), \qquad g \quad G_1, x \quad \mathbf{L}(G_1).$$

Hint: Consider $\exp_{G_2}(t\mathbf{L}(\phi)(gxg^{-1}))$ and use that $\exp_{G_2}$ is injective in a 0-neighborhood.

2. Use the uniqueness in Proposition 1.8.1 to prove Lemma 1.8.3, i.e. $\mathbf{L}(\mathrm{id}_G) = \mathrm{id}_{\mathbf{L}(G)}$ and $\mathbf{L}(\phi \circ \psi) = \mathbf{L}(\phi) \circ \mathbf{L}(\psi)$.

3. Let $G \quad \mathrm{GL}_n(\mathbb{K})$ be a linear Lie group. For $g \quad G$ we define the conjugate automorphism $c_g \colon G \to G, c_g(h) = ghg^{-1}$.

a) Prove that c_g is a continuous group homomorphism for every $g \quad G$.

b) Prove that $\mathbf{L}(c_g) = \mathrm{Ad}(g)$ for every $g \quad G$, i.e. the adjoint representation from Lemma 1.5.4 is given by derivatives of the conjugation action.

c) Prove that $\mathrm{Ad} \colon G \to \mathrm{GL}(\mathbf{L}(G)), g \to \mathrm{Ad}(g)$ is a continuous group morphism. Deduce that the adjoint representation is a continuous representation of G by Lie algebra homomorphisms $\mathrm{Ad}(g)$ on $\mathbf{L}(G)$.

d) Define $\mathrm{ad}(x) \colon \mathbf{L}(G) \to \mathbf{L}(G), \mathrm{ad}(x)(y) = [x, y]$ for $X \quad \mathbf{L}(G)$. Show that $\mathbf{L}(\mathrm{Ad}) = \mathrm{ad}$. In other words, the derived representation of the adjoint representation is for each x given by the Lie bracket.
Hint: Review the proof of Proposition 1.5.9 and (1.12).

e) Deduce that for every $X \quad M_n(\mathbb{K})$ the following holds

$$\mathrm{Ad}(\exp(X)) = \exp(\mathrm{ad}(X)). \tag{1.13}$$

4. Verify some details in Example 1.8.9. In particular:

a) Compute explicitely some matrix products in the normaliser $N(T)$.

b) Show that for any $M, N \quad T$ there is $H \quad T$ such that $HMH^{-1} = N$.

c) Convince yourself that the group is indeed highly non-commutative as claimed.

1.9. The Baker-Campbell-Dynkin-Hausdorff Formula

In this section, we derive a formula which expresses $\exp(x)\exp(y)$ as the exponential image $\exp(x \ast y)$ of an element $x \ast y$ which can be described in terms of iterated commutator brackets. This implies in particular, that the group multiplication in a I_n-neighborhood of $\mathrm{GL}_n(\mathbb{K})$ is completely determined by the commutator bracket $[\cdot,\cdot]$. As a consequence the Lie algebra of a linear Lie group G at least locally completely determines G.

The local multiplication "$\ast$" is called the *Baker-Campbell-Dynkin-Hausdorff (BCDH) multiplication*[12] and the following identity *BCDH-formula*

$$x \ast y := \log(\exp(x)\exp(y)) = x + y + \frac{1}{2}[x,y] + \frac{1}{12}[x,[x,y]] + \frac{1}{12}[y,[y,x]] + \cdots. \tag{1.14}$$

Note that for matrices $XY = YX$ we have $[X,Y] = 0$ and thus (1.14) becomes for commuting matrices just the statement $e^{X+Y} = e^X e^Y$ from Lemma 1.4.3.

The idea to establish the BCDH-formula is to use a similar trick as we did in the construction of the Lie algebra, where we proved, (1.8), that

$$[X,Y] = \frac{d}{dt}\Big|_{t=0} e^{tX} Y e^{-tX}.$$

Let us illustrate the idea by establishing the following.

1.9.1 Proposition (Toy BCDH-formula) *Let $X, Y \in M_n(\mathbb{K})$ be a pair of matrices with*

$$[X,[X,Y]] = 0, \qquad [Y,[X,Y]] = 0.$$

Then

$$e^X \cdot e^Y = e^{X+Y+\frac{1}{2}[X,Y]}. \tag{1.15}$$

1.9.2 Some preparation concerning notation is needed. Let $X \in M_n(\mathbb{K})$, then we let $\lambda_X, \rho_X : M_n(\mathbb{K}) \to M_n(\mathbb{K}), \lambda_X(Y) = XY, \rho_X(Y) = YX$ be the left /right multiplication with X and $\mathrm{ad}(X) = \lambda_X - \rho_X$. Note $\mathrm{ad}(X)(Y) = [X,Y]$.

Proof of Proposition 1.9.1. Write $Z = [X,Y]$ and note that Z commutes both with X and Y. Then define $\gamma(t) = e^{tX}e^{tY}e^{-t^2Z/2}$, $t \in \mathbb{R}$. Compute $\gamma'(t)$ using Exercise 1.4 4:

$$\begin{aligned}\gamma'(t) &= e^{tX}Xe^{tY}e^{-t^2Z/2} + e^{tX}e^{tY}Ye^{-t^2Z/2} - e^{tX}e^{tY}tZe^{-t^2Z/2} \\ &= e^{tX}e^{tY}(e^{-tY}Xe^{tY} + Y - tZ)e^{-t^2Z/2}\end{aligned}$$

[12] In most of the literature BCDH is just called BCH (removing Dynkin from the acronym). We follow [HN12] in the naming convention. Baker, Campbell and Hausdorff stated the formula in its original quantitative form, while Dynkin (later!) gave the numerical values for the coreefficients, cf. https://en.wikipedia.org/wiki/Baker-Campbell-Hausdorff_formula.

We compute now with Exercise 1.8 3 e) an expression for the first term in the bracket

$$e^{-tY}Xe^{tY} = \mathrm{Ad}(e^{-tY})(X) = e^{-t\,\mathrm{ad}(Y)}X = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\mathrm{ad}(Y)^kX = X - t[Y,X] = X + tZ$$

Here we have used that $\mathrm{ad}(Y)^0(X) = X$ and higher powers of $\mathrm{ad}(Y)$ applied to X vanish (as the iterated Lie brackets are 0). We insert this in the above calculation and get

$$\gamma'(t) = e^{tX}e^{tY}(X + tZ + Y - tZ)e^{-t^2Z/2} = \gamma(t)(X+Y). \tag{1.16}$$

The last equality follows from Exercise 1.4 6.: As Z commutes with $X+Y$ (it commutes with both X and Y), $e^{-t^2Z/2}$ commutes with $X+Y$ and the three factors then form $\gamma(t)$. As $\gamma(0) = I_n$, (1.16) means that γ is a 1-parameter group with derivative $X+Y$ in $t=0$. Whence by the 1-parameter group Theorem 1.4.11 we obtain

$$e^{t(X+Y)} = \gamma(t) = e^{tX}e^{tY}e^{-t^2Z/2}.$$

As Z commutes with $X+Y$ evaluating at $t=1$ yields then (1.15). □

The proof of Proposition 1.9.1 gives a hint at what one would need to do to establish the full BCDH formula. The idea is to compute an expression for the derivative of exp:

1.9.3 Proposition *Let $X \in M_n(\mathbb{K})$. Then*

$$\mathbf{d}\exp(X) = \lambda_{\exp(X)} \circ \int_0^1 e^{-s\,\mathrm{ad}(X)}\,\mathrm{d}s = \lambda_{\exp(X)} \circ \sum_{n=0}^{\infty}(-1)^n\frac{\mathrm{ad}(X)^m}{(m+1)!} \tag{1.17}$$

Now we need to find a formula for $\mathbf{d}\exp(X)^{-1}$ (since this gives the derivative of log). For this one writes exp in (1.17) as a power series, integrates termwise and searches for a series which mutliplied with the resulting series gives 1. It turns out that

$$\Psi(z) = \frac{z\log(z)}{z-1} := z\sum_{k=0}^{\infty}\frac{(-1)^k}{k+1}(z-1)^k \quad \text{for } |z-1| < 1$$

does the job, but we are lazy and will not go into all the work here (in case you are interested take a look at [HN12, Section 3.4]). We just state the result without a proof

1.9.4 Theorem (BCDH-formula) *For $X, Y \in M_n(\mathbb{K})$ with $\|X\|, \|Y\|$ sufficiently small the integral formulation of the BCDH-formula holds*

$$X * Y = \log\left(e^Xe^Y\right) = X + \int_0^1 \Psi(e^{\mathrm{ad}(X)}e^{\mathrm{ad}(tY)})Y\,\mathrm{d}t.$$

Inserting the Taylor expansion for Ψ and integrating termwise one obtains the series formulation of the BCHD-formula

$$X * Y = X + \sum_{\substack{k,m\geq 0\\ p_i+q_i>0}} \frac{(-1)^k \operatorname{ad}(X)^{p_1}\operatorname{ad}(Y)^{q_1}\cdots \operatorname{ad}(X)^{p_k}\operatorname{ad}(Y)^{q_k}\operatorname{ad}(X)^m}{(k+1)(q_1+\cdots+q_k+1)p_1!q_1!\cdots p_k!q_k!m!}Y. \tag{1.18}$$

Collecting all terms of low order and rewriting the nested ad*-operators as Lie brackets, one obtains* (1.14).

We have not quantified in Theorem 1.9.4 what it means that the matrices are small enough. One can show that (1.18) converges if $\|X\| + \|Y\| < \ln(2)/2$. Let us however give an example of matrices for which (1.18) does not converge.

1.9.5 Example Consider the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ of all complex 2×2-matrices with trace 0 (cf. Example 1.6.4). For

$$X := \begin{pmatrix} 0 & i\pi \\ i\pi & 0 \end{pmatrix}, \qquad Y := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

in $\mathfrak{sl}_2(\mathbb{C})$ one calculates (cf. Remark 1.4.8 for a similar calculation)

$$e^X e^Y = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix} \tag{1.19}$$

Now assume that $Z \in \mathfrak{sl}_2(\mathbb{C})$ such that $e^Z = e^X e^Y$. If Z has distinct eigenvalues, then Z will be diagonalizable, and so the right hand side of (1.19) must be diagonalizable by Exercise 1.4.2 (which it is not!). We conclude that Z must have a repeated eigenvalue. Since $\mathrm{tr}Z = 0$ the repeated eigenvalue must be 0 (the trace is the sum of eigenvalues of a matrix with multiplicities, see [Mey23, Section 7.1]). We thus find $0 \neq v \in \mathbb{C}^2$ with $Zv = 0$ from which it follows that $e^Z v = e^0 v = v$ and e^Z has 1 as eigenvalue. However, the right hand side of (1.19) does not have the eigenvalue 1. In conclusion there can not be a $Z \in \mathfrak{sl}_2(\mathbb{C})$ such that $e^Z = e^X e^Y$. Thus $X * Y$ can not converge in $\mathfrak{sl}_2(\mathbb{C})$. Further, $\exp_{\mathrm{SL}_2(\mathbb{C})}$ is not surjective onto $\mathrm{SL}_2(\mathbb{C})$.

1.9.6 Remark

- The BCDH-formula only needs a Lie bracket to define a multiplication in terms of $\operatorname{ad}(X) = [X, \cdot]$. Then the proof of Theorem 1.9.4 applies and yields an open neighborhood on which the BCDH-multiplication is defined. For a linear Lie group, this local multiplication encodes the multiplication of Lie group.
- Let $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism, then for all $x, y \in \mathfrak{g}$ we find that

$$\varphi(X * Y) = \varphi(X) * \varphi(Y).$$

1.9.7 Definition (Local Lie group) A *local Lie group* is an open 0-neighborhood $U \subseteq E$ in a finite dimensional vector space E such that there are smooth maps $\mu\colon U \times U \to E$ (*local multiplication*) and $\iota\colon U \to E$ (*inversion*), such that, for X, Y, Z sufficiently close to 0 in E the following holds

$$\mu(X, 0) = X = \mu(0, X) \quad \mu(X, \iota(X)) = 0 = \mu(\iota(X), X)$$
$$\mu(X, \mu(Y, Z)) = \mu(\mu(X, Y), Z)$$

Let $(U_i, \mu_i, \iota_i), i = 1, 2$ be local Lie groups, a *morphism of local Lie groups* is a smooth map $\varphi\colon V \to U_2$ defined on an open 0-neighborhood such that there exists $0 \in W \overset{\circ}{\subseteq} V$ with $\mu_1(X, Y), \iota_1(X) \in V$ if $X, Y \in W$ and

$$\varphi(\mu_1(X, Y)) = \mu_2(\varphi(X), \varphi(Y)), \quad \varphi(\iota_1(X)) = \iota_2(\varphi(X)).$$

If φ can be choosen as a diffeomorphism onto an open 0-neighborhood in U_2, we say that the local Lie groups are isomorphic.

Here is a last result about the BCDH-formula we will not prove[13]

1.9.8 Lemma (Local associativity of the BCHD-muliplication) *Let* $(\mathfrak{g}, [\cdot, \cdot])$ *be a finite dimensional Lie algebra. Then the BCHD-series* $X * Y$ *defined via* (1.18) *converges for* $X, Y, Z \in \mathfrak{g}$ *sufficiently near* 0*. On its domain of definition it yields a smooth map such that the following holds*

$$(X * Y) * Z = X * (Y * Z)$$

as long as both sides are defined.

We can now construct from every finite dimensional Lie algebra a local Lie group:

1.9.9 Theorem (Lie algebras integrate to local Lie groups) *Every finite dimensional Lie algebra* $\mathfrak{g}$ *admits an open* 0*-neighborhood* $U_\mathfrak{g}$ *which becomes a local Lie group with respect to the mappings* $\mu(x, y) = x * y$ *defined via the BCDH-formula and* $\iota(x) = -x$*.*

Further, every Lie algebra morphismn $\psi\colon \mathfrak{g} \to \mathfrak{h}$ *gives rise to a morphism of local Lie groups* $U_\mathfrak{g} \to U_\mathfrak{h}$

This is only the tip of the iceberg. For deeper results cf. Section 3.4, but before we need to step away from matrices and consider Lie groups as groups which are manifolds.

[13] The proof turns out to be either a long calculation involving power series or a technical argument we do not wish to go into. We recommend Terry Tao's blog for a proof https://terrytao.wordpress.com/2011/10/29/associativity-of-the-baker-campbell-hausdorff-formula/.

1.9 Exercises

1. We consider the Heisenberg algebra $\mathfrak{h} = \operatorname{span}(X, Y, Z) \subset M_3(\mathbb{R})$, where

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

a) Show that $\mathfrak{h}$ is a Lie subalgebra of $M_3(\mathbb{R})$.

b) Find an explicit formula for $\exp(aX + bY + cZ)$.

c) Show that the matrix exponential exp restricts to an invertible continuous map $\mathfrak{h} \to \exp(\mathfrak{h}) \subset \mathrm{GL}_3(\mathbb{R})$ with continuous inverse.

d) Use (b) and Proposition 1.9.3, respectively, to compute $\frac{d}{dt}\Big|_{t=0} e^{A+tB}$ when $A, B \in \mathfrak{h}$ in two different ways.

e) Construct a map $\mu\colon \mathfrak{h} \times \mathfrak{h} \to \mathfrak{h}$ such that

$$e^A \cdot e^B = e^{\mu(A,B)}, \quad \forall A, B \in \mathfrak{h}.$$

Is $(\mathfrak{h}, \mu)$ a group? I.e. can we say that the local multiplication of the Heisenberg algebra gives rise to a Lie group?
Remark: It is instructive to compare this to Theorem 1.9.9.

2. Prove Theorem 1.9.9 (use Lemma 1.9.8 without proof or see Footnote 13).

3. Use the BCDH-formula to give an alternative proof of the Trotter product formula Proposition 1.4.13.
Hint: Let $X, Y \in M_n(\mathbb{K})$. Plug $k(X/k) \oplus (Y/k)$ for $k \in \mathbb{N}$ into the formula (1.14) and study the limit $k \to \infty$.

4. Let $U = B_{1/2}(0) := \{x \in \mathbb{R} : |x| < 1/2\}$ and

$$\mu(x, y) = \frac{2xy - x - y}{xy - 1}, \quad \iota(x) = \frac{x}{2x - 1}.$$

a) Show that this multiplication and inversion turns U into a (commutative) local Lie group.

b) Prove that $\kappa\colon B_1(0) \to \mathbb{R}$ with $\kappa(x) = \frac{x}{x-1}$ satisfies $\kappa(x+y) = \mu(\kappa(x), \kappa(y))$ and $\kappa(-x) = \iota(\kappa(x))$, where defined. This implies that up to the map κ the local multiplication and inversion in U is just the usual addition and inversion of the real numbers $\mathbb{R}$.

2. Manifolds, vector fields and flows

Motivation: Why Lie groups and manifolds?

From now on, we step outside of the matrix world and use the language of manifolds to construct an abstract setting for Lie theory. It gets cumbersome to study structure of Lie groups while going back all the time to a realisation as a group of matrices. For example the unit circle

$$S^1 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

carries a nice group structure by identifying $\mathbb{R}^2 = \mathbb{C}$ and using the multiplication of complex numbers. In view of $S^1 = \{e^{it} : t \in [0, 2\pi]\}$ the multiplication should be smooth and the circle should be a Lie group. It is not hard to identify the circle as a matrix Lie group (cf. Example 1.8.9) but it would be easier if we could be working directly with S^1. In particular, it would be nice to get rid of the somewhat artificial requirement that the mappings are defined on some ambient space (like $\mathrm{GL}_n(\mathbb{K})$) and smooth there.

The other reason is that we can then formulate Lie group actions not directly related to linear structures. In the introduction we considered symmetries of differential equations via 1-parameter groups. The following example is developed in [Olv93, Example 2.43]:

2.0.1 Example (Wave equation) Consider the wave equation in 2d

$$u_{tt} - u_{xx} - u_{yy} = 0$$

Here subscripts mean partial derivatives, u depends on time t and position x, y. So the solution curves $(x, y, t, u(t, x))$ lie in $\mathbb{R}^4$. One can work out the symmetries and with $f_\varepsilon(x, y, t) := \sqrt{1 - 2\varepsilon x - \varepsilon^2(t^2 - x^2 - y^2)}$ a 1-parameter group of symmetries is

$$\Gamma_\varepsilon(x, y, t, u) = \left(\frac{x + \varepsilon(t^2 - x^2 - y^2)}{f_\varepsilon(x, y, t)^2}, \frac{y}{f_\varepsilon(x, y, t)^2}, \frac{t}{f_\varepsilon(x, y, t)^2}, f_\varepsilon(x, y, t)u\right).$$

So we maybe could express this group as a matrix group it seems impractical at best.

Finally, we will be interested in actions of Lie groups on manifolds. When matrices act on $\mathbb{R}^n$ by matrix vector multiplication, this will turn out to be an action of a (linear) Lie group. As we shall see, each such action of a Lie group yields a group morphism into the diffeomorphism group of the manifold. These diffeomorphism groups turn out to be Lie groups. Alas, they are infinite-dimensional Lie groups whence beyond the course content (see e.g. [Sch23] for more information).

2.1. Manifolds

In this section we discuss the basic notions of manifolds. Conceptually, a manifold is some set with additional structure which allows us to use the tools of calculus. As we learn in multivariable calculus (repeated in Appendix A.2) we can do calculus on open subsets of $\mathbb{R}^n$. The subsets we have in mind will usually not be (directly) identifiable as open subsets. Instead they will be some sort of curved space such as the unit sphere. The idea is then to use special maps, the coordinate charts, which allow us to transport from the curved space to the familiar setting of open subsets of euclidean space.

2.1.1 Definition (Charts and atlas) Let M be a Hausdorff topological space and fix $d \in \mathbb{N}_0$. A *chart* for M is a homeomorphism $\phi\colon U_\phi \to V_\phi$ from $U_\phi \subseteq M$ onto $V_\phi \subseteq \mathbb{R}^d$. Let $r \in \mathbb{N}_0 \cup \{\infty\}$. A C^r*-atlas* for M is a set $\mathcal{A}$ of charts for M satisfying the following

(a) $M = \bigcup_{\phi \in \mathcal{A}} U_\phi$

(b) for all $\phi, \psi \in \mathcal{A}$ the *change of charts* $\phi \circ \psi^{-1}$ (which are mappings between open subsets of $\mathbb{R}^d$) are C^r.[1]

Two C^r-atlases $\mathcal{A}, \mathcal{A}'$ for M are *equivalent* if their union $\mathcal{A} \cup \mathcal{A}'$ is a C^r-atlas for M. This is an equivalence relation.

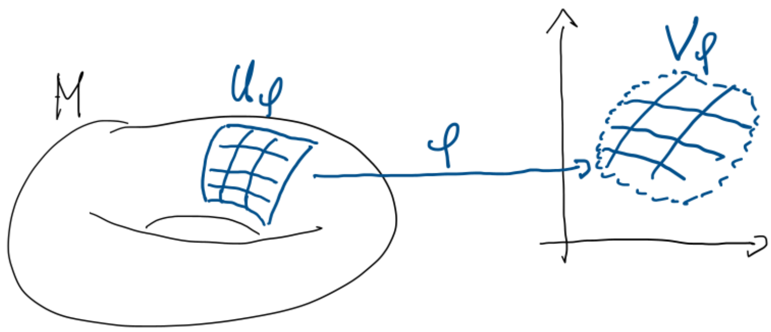


Figure 2.1.: Illustration of a chart (U_ϕ, ϕ) on the torus M mapping the chart domain U_ϕ to an open subset of the model space $\mathbb{R}^2$ of the manifold.

2.1.2 Remark For $M = \mathbb{R}$, the maps $(M, \mathrm{id}_\mathbb{R})$ and (M, ψ) with $\psi(x) = x^3$ are homeomorphisms, whence charts for M as a manifold. These charts are not C^1-compatible as the map $\psi \circ \mathrm{id}_\mathbb{R} = \psi$ is smooth, but its inverse is not differentiable.

[1] Note that the change of charts $\phi \circ \psi^{-1}$ is always a homeomorphism (between the sets on which it is defined), since charts are homeomorphisms. That the change of charts map is a C^r-diffeomorphism is an additional requirement. Formally, if the charts do not intersect, the change of charts is the empty map $\emptyset \to \emptyset$ which is C^r.

2.1.3 Definition A C^r manifold $(M, \mathcal{A})$ is a Hausdorff topological space with an equivalence class of C^r-atlases $\mathcal{A}$. (If the equivalence class $\mathcal{A}$ is clear we simply write M.) We say that M is a *d-dimensional manifold* if it admits a manifold atlas of charts mapping to open sets of $\mathbb{R}^d$. These charts will also be called *d-charts* to stress the dimension.

Examples of manifolds

2.1.4 Example Every $\mathbb{R}^d$ is a manifold with global chart given by the identity $\mathrm{id}_{\mathbb{R}^d}$. Similarly, every $U \overset{\circ}{\subseteq} \mathbb{R}^d$ is a manifold with global chart given by the inclusion $U \to \mathbb{R}^d$.

2.1.5 Example (Unit sphere in $\mathbb{R}^d$) For $d \in \mathbb{N}$ we endow $\mathbb{R}^d$ with the euclidean inner product $\langle \cdot, \cdot \rangle$. The *unit sphere* $S^{d-1} := \{x \in \mathbb{R}^d \mid \langle x, x \rangle = 1\}$ is a C^∞-manifold. To construct charts, we consider the subspace $H_{x_0} = \{y \in \mathbb{R}^d \mid \langle y, x_0 \rangle = 0\} \subseteq \mathbb{R}^d$ for $x_0 \in S^{d-1}$. Note that H_{x_0} is isomorphic to $\mathbb{R}^{d-1}$. Define $U_{x_0} := \{x \in S^{d-1} \mid \langle x, x_0 \rangle > 0\}$ and $V_{x_0} := \{y \in H_{x_0} \mid \langle y, y \rangle < 1\}$. We obtain a chart centered at x_0 via

$$\phi_{x_0} \colon U_{x_0} \to V_{x_0}, \quad x \mapsto x - \langle x, x_0 \rangle x_0.$$

(its inverse is given by the formula $\phi_{x_0}^{-1}(y) = y + \sqrt{1 - \langle y, y \rangle}\, x_0$). Applying these formulae, we see that the change of charts map for $x_0, z_0 \in S^{d-1}$ is a smooth map between open (possibly empty) subsets of euclidean space:

$$\phi_{z_0} \circ \phi_{x_0}^{-1}(y) = (y - \langle y, z_0 \rangle z_0) + \sqrt{1 - \langle y, y \rangle}\,(x_0 - \langle x_0, z_0 \rangle z_0).$$

We will usually only consider manifolds with the Hausdorff property (i.e. if $x, y \in M$ are distinct points, then there are open disjoint sets U, V such that $x \in U$ and $y \in V$). However, here is an example of a non-Hausdorff smooth manifold:

2.1.6 Example (The line with the two origins) Consider the set $M = \{1\} \times \mathbb{R} \cup \{2\} \times \mathbb{R}/\sim$, where we identify $(1, x) \sim (2, x)$ if $x \neq 0$. We put on M the final topology with respect to the two maps

$$\phi_i^{-1} \colon \mathbb{R} \to M, x \mapsto [(i, x)], \quad i = 1, 2.$$

Then $\phi_i = (\phi_i^{-1})^{-1}, i = 1, 2$ is a homeomorphism from an open set onto $\mathbb{R}$ and the two maps form a smooth atlas for M (Exercise!). Note that $[(1, 0)] = [(2, 0)]$ but the two points have no disjoint open neighborhoods. So M is not Hausdorff. Non-Hausdorff manifold are amusing but not of interest to us as limits are no longer unique. So we would have problems to define differential calculus on them.

2.1.7 (Only Hausdorff manifolds!) From now on: all manifolds will be Hausdorff.

2.1.8 Definition Let M be a C^r-manifold and together with vector subspace $F \subseteq \mathbb{R}^d$. A *(C^r-)submanifold of M* is a subset $N \subseteq M$ such that for each $x \in N$, there exists a chart $\varphi\colon U_\varphi \to V_\varphi$ of M around x such that $\varphi(U_\varphi \cap N) = V_\varphi \cap F$. Then $\varphi_N := \varphi|_{U_\varphi \cap N}^{V_\varphi \cap F}$ is a chart for N, called a *submanifold chart*. The submanifold charts form a C^r-atlas for N. If $\dim F = k$, we say that N is a *k-dimensional submanifold* of M.

If $\dim M = n$ and $\dim N = n - 1$ we also call N a *hypersurface* in M.

2.1.9 Example (The trivial example) $U \subseteq^\circ \mathbb{R}^d$ is a submanifold of $\mathbb{R}^d$. Namely, the global chart $(\mathrm{id}_{\mathbb{R}^d}, \mathbb{R}^d)$ of $\mathbb{R}^d$ restricts to a submanifold chart

$$\iota_U := \mathrm{id}_{\mathbb{R}^d}|_U^U\colon U \to U \subseteq^\circ \mathbb{R}^d.$$

We shall now prove that the unit spheres $\mathbb{S}^{d-1}$ are also submanifolds. For this we prove a handy criterion to construct submanifolds via the solution sets of non-linear equations:

2.1.10 Definition Let $f\colon U \to \mathbb{R}^m$ be a C^1-map. We call $y \in \mathbb{R}^m$ a *regular value of f* if for each $x \in U$ with $f(x) = y$ the differential $\mathbf{d}f(x)$ is surjective. Otherwise y is called a *singular value of f*. Note that in particular, each $y \in \mathbb{R}^m \setminus f(U)$ is a regular value.

2.1.11 Proposition (Regular value theorem - local version) *Let $U \subseteq^\circ \mathbb{R}^n$, $f\colon U \to \mathbb{R}^m$ a smooth map and $y \in \mathbb{R}^m$ a regular value of f. Then $M := f^{-1}(y)$ is an $(n-m)$-dimensional submanifold of $\mathbb{R}^n$, hence in particular a smooth manifold.*

Proof. In this proof we use subscripts for maps and points to denote components. Let $d := n - m$ and observe that $d > 0$ as $\mathbf{d}f(x)\colon \mathbb{R}^n \to \mathbb{R}^m$ is surjective for each $x \in M$. We have to construct for each $x_0 \in M$ an x_0-neighborhood $V \subseteq^\circ \mathbb{R}^n$ and a diffeomorphism

$$\phi\colon V \to \phi(V) \subseteq \mathbb{R}^n, \text{ with } \phi(V \cap M) = \phi(V) \cap (\mathbb{R}^d \times \{0\}).$$

After permuting the coordinates, we may assume that the vectors

$$\mathbf{d}f(x_0)(e_{d+1}), \ldots, \mathbf{d}f(x_0)(e_n)$$

form a basis of $\mathbb{R}^m$. Set $f(x_0) = (y_1, \ldots, y_m) \in \mathbb{R}^m$, then we consider the map

$$\phi\colon U \to \mathbb{R}^n, x = (x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_d, f_1(x) - y_1, \ldots, f_m(x) - y_m).$$

The first components of ϕ are just the projection onto the first d elements. Hence

$$\mathbf{d}\phi(x_0)(e_j) = \begin{cases} (e_j, \mathbf{d}f(x_0)(e_j)) & \text{for } j \leq d, \\ (0, \mathbf{d}f(x_0)(e_j)) & \text{for } j > d, \end{cases}$$

where the first element in the above pair is a vector in $\mathbb{R}^d$ and the second a vector in $\mathbb{R}^m$. It follows that the linear map $\mathbf{d}\phi(x_0)\colon \mathbb{R}^n \to \mathbb{R}^n$ is invertible. Hence by the Inverse function Theorem A.2.7 there exists an open x_0-neighborhood $V \subseteq U$ for which $\phi|_V : V \to \phi(V)$ is a diffeomorphism onto an open subset of $\mathbb{R}^n$. Since

$$M = \{p \in U\colon \phi(p) = (\phi_1(p), \ldots, \phi_d(p), 0, \ldots, 0)\} = \phi^{-1}(\mathbb{R}^d \times \{0\}),$$

we have $\phi(M \cap V) = \phi(V) \cap (\mathbb{R}^d \times \{0\})$ and ϕ is a submanifold chart around x_0. □

2.1.12 Example The regular value theorem is easy to apply for hypersurfaces, i.e. if $m = 1$. Then $f\colon U \to \mathbb{R}$ is a smooth function and the condition that $\mathbf{d}f(X)$ is surjective simply means that $\mathbf{d}f(x) \neq 0$, i.e. there is a partial derivative which is non-zero at x.

Let $A = (a_{ij}) \in M_n(\mathbb{R})$ be a symmetric matrix and

$$f(x) := x^\top Ax = \sum_{i,j=1}^{n} a_{ij}x_ix_j.$$

The corresponding *quadric* $Q := \{x \in \mathbb{R}^n\colon f(x) = 1\}$ is a submanifold of $\mathbb{R}^n$. To see this, assume that $x \in \mathbb{R}^n$ with $f(x) = 1$. Then

$$\mathbf{d}f(x)v = v^\top Ax + x^\top Av = 2v^\top Ax.$$

Therefore $\mathbf{d}f(x) = 0$ is equivalent to $Ax = 0$, which is never the case if $x^\top Ax = 1$. Similarly, all level surfaces of f are smooth hypersurfaces of $\mathbb{R}^n$.

For $A = I_n$ we have $Q = \mathbb{S}^{n-1}$, the $(n-1)$-dimensional unit sphere. For $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and nonzero λ_i we obtain the *hyperboloids*

$$Q = \left\{ x \in \mathbb{R}^n\colon \sum_{i=1}^{n} \lambda_i x_i^2 = 1 \right\}$$

For singular values the level sets may or may not be submanifolds (Exercise).

2.1.13 Definition Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be C^r manifolds. Then the product $M \times N$ becomes a C^r-manifold using the atlas $\mathcal{C} := \{\varphi \times \psi \mid \varphi \in \mathcal{A}, \psi \in \mathcal{B}\}$. We call the resulting C^r manifold the *(direct) product of M and N*.

2.1.14 Definition Let $r \in \mathbb{N}_0 \cup \{\infty\}$ and M, N be C^r manifolds. A map $f\colon M \to N$ is called C^r if f is continuous and, for every pair of charts φ, ψ, the map

$$\psi \circ f \circ \varphi^{-1}\colon \mathbb{R}^n \supseteq \varphi(f^{-1}(U_\psi) \cap U_\varphi) \to \mathbb{R}^m$$

is a C^r map. We write $C^r(M, N)$ for the set of all C^r-maps from M to N.

A smooth map $f: M \to N$ is called a *diffeomorphism*, if there exists a smooth map $g: N \to M$ with $g \circ g = \mathrm{id}_n$ and $g \circ f = \mathrm{id}_M$. We also write $\mathrm{Diff}(M)$ for the set of all diffeomorphisms of M.

If $I \subseteq \mathbb{R}$ is an open interval, then a smooth map $\gamma: I \to M$ is called a smooth curve. For a not necessarily open interval $I \subseteq \mathbb{R}$ we call $\gamma: I \to \mathbb{R}^n$ smooth if all derivatives $\frac{d^k}{dt^k}\gamma$ exist and define continuous functions $I \to \mathbb{R}^n$. Hence such a curve with values in M is called smooth if for every chart (ϕ, U) of M we have thart $\phi \circ \gamma: \gamma^{-1}(U) \to \phi(U)$ is a smooth curve.

2.1.15 Remark Let $f: M \to N$ be a continuous map between C^r-manifolds. Assume that for some charts (U_ϕ, ϕ) and (U_ψ, ψ) the composition $\phi \circ f \circ \psi$ is $C^r, r \in \mathbb{N} \cup \{\infty\}$. Then for any other pair of charts (U_κ, κ) and (U_λ, λ) with $f(U_\lambda \cap U_\psi) \subseteq U_\phi \cap U_\kappa$ we have on $U_\lambda \cap U_\psi$ that

$$\kappa \circ f \circ \lambda^{-1}|_{\lambda(U_\lambda \cap U_\psi} = (\kappa \circ \phi^{-1}) \circ (\phi \circ f \circ \psi^{-1}) \circ (\psi \circ \lambda^{-1}|_{\lambda(U_\lambda \cap U_\psi)})$$

where the mapping in the middle is C^r by assumption and the other mappings are change of charts (which are C^r by M, N being C^r-manifolds). Hence $\kappa \circ f \circ \lambda^{-1}|_{\lambda(U_\lambda \cap U_\psi)}$ is also C^r by the chain rule. This argument is called "insertion of charts" and we leave it from now on to the reader. With the insertion of charts argument, it is easy to see that:

(a) it suffices to test the C^r-property with respect to any atlas of M and N.

(b) if $f: M \to N$ and $g: N \to L$ are C^r-maps, so is $g \circ f: M \to L$.

(c) If M, N_1, N_2 are C^r-manifolds and $f_i: M \to N_i, i = 1, 2$ are mappings. Then $f := (f_1, f_2): M \to N_1 \times N_2$ is C^r if and only if f_1, f_2 are C^r

2.1.16 Lemma *Let N be a (C^r-)submanifold of the C^r-manifold M. Then the inclusion $\iota: N \to M$ is C^r. Further, $f: P \to N$ is C^r if and only if $\iota \circ f$ is C^r.*

Proof. Due to Remark 2.1.15 if suffices to check the C^r-property of ι in charts which cover N. Choose charts (φ, U_φ) of M which induce submanifold charts φ_N, Definition 2.1.8. Then $\varphi \circ \iota \circ \varphi_N^{-1}$ is the inclusion $V_\varphi: F \to V$ which is C^r.

If f is C^r, so is $\iota \circ f$ by Remark 2.1.15. Conversely, let $\iota \circ f$ be a C^r-map and φ, φ_N as before and ψ a chart for P. Then $\varphi \circ \iota \circ f \circ \psi^{-1}: \psi(U_\psi \cap f^{-1}(U_\varphi) \to \mathbb{R}^d$ is C^r with values in the subspace F and thus $(\varphi \circ \iota \circ f \circ \psi)|^F = \varphi_N \circ f \circ \psi^{-1}$ is C^r. We conclude that f is C^r. □

In these notes we give only a short introduction to the theory of manifolds and provide the essentials for the treatment of Lie groups. More information is contained in [HN12] or the more extensive sources such as [Lee13],

2.1 Exercises

1. Let E be a vector space with base $B = (b_1, \dots, b_n)$, then $\phi_B : E \to \mathsf{R}^d$, $\phi_B(b_i) = e_i$ (ith unit vector in R^n) is a vector space isomorphism. Show that if we endow E with a topology making ϕ_B a homeomorphism, then all mappings ϕ_B for different choices of bases B become homeomorphisms and define a smooth atlas turning E into an n-dimensional manifold.

2. Verify the details of Example 2.1.5: Check that the charts make sense as mappings from U_{x_0} to V_{x_0}. Show that the change of charts $\phi_{x_0} \circ \phi_{z_0}^{-1}$ is smooth for all x_0, z_0 S^{d-1} such that $U_{x_0} \cap U_{z_0} =$.

3. Verify all the details in Example 2.1.6.

4. We investigate level sets for singular values of smooth functions. w

a) For $f : \mathsf{R}^2 \to \mathsf{R}, \quad f(x_1, x_2) = x_1 x_2$ check that 0 is a singular value. Then show that the preimage $f^{-1}(0)$ with the subspace topology is not a submanifold of R^2 (draw it!).

b) For $g : \mathsf{R}^2 \to \mathsf{R}, \quad g(x_1, x_2) = x_1^2 + x_2^2$ check that 0 is a singular value. Prove that $g^{-1}(0)$ with the subspace topology is a 0-dimensional submanifold of R^2.

5. By Example 2.1.12, S^1 is a submanifold of R^2 by applying the local regular value theorem (i.e. Proposition 2.1.11) to $f^{-1}(1)$ for the map $f(x, y) = x^2 + y^2$. Revisit the proof of the local regular value theorem and describe explicitely the charts around x_0 constructed there for the example S^1.
Hint: Consider first $x_0 = (a, b)$ S^1 with $b = 0$. Can you compute ϕ^{-1} in this case?

6. Familiarize yourself with the insertion of charts argument and prove (a)-(c) in Remark 2.1.15.

7. Let $f : M \to N$ be a C^r-map between C^r-manifolds. show that the graph graph$(f) := \{(m, f(m)) \mid m \quad M\}$ is a submanifold of $M \times N$.
Hint: Use the description of the graph to construct submanifold charts by hand.

8. Let M be a manifold and $U \circ M$. Let A be an atlas of M. Endow U with the subspace topology and show that $A_U := \{\varphi|_{U \cap U_\varphi} \mid (\varphi, U_\varphi) \quad A\}$ is a manifold atlas for U turning it into a submanifold of M.

9. Check that the set C in Definition 2.1.13 is a C^r atlas for the product manifold.

2.2. Tangent spaces and the tangent bundle

The advantage of smooth manifolds is that they permit us to analyze differentiable structures and obtain a notion of smoothness for maps between manifolds. While we have defined maps to be differentiable when they are differentiable after composition with charts, we are still lacking a replacement of the derivative of a differentiable map. So if $f \colon M \to N$ is differentiable, M, N (smooth) manifolds, we lack a replacement for

$$\mathbf{d}f(x)(h) = \lim_{t\to 0} \frac{1}{t}(f(x+th) - f(x)). \tag{2.1}$$

The problem is that nothing in (2.1) makes sense on manifolds! Since the manifold M does not have an addition, $x + th$ is not even defined. Similarly, multiplication with $\frac{1}{t}$ and the expression $f(x + th) - f(x)$ make no sense on N. To see by what we should replace this, recall that Appendix A.2 states that $\mathbf{d}f(p)(h)$ should show how f changes at p in the direction h. Directions on the manifold can be tracked using curves. Hence the idea is to replace $x + th$ (ill defined on M) by a curve γ whose derivative at p is h. Our replacement will then have to transform these directions to something on N.

2.2.1 Definition Let M be a smooth manifold and $p \quad M$. We say that a C^1-curve γ *passes through* p if $\gamma(0) = p$. For two such curves γ, η we define the relation

$$\gamma \quad \eta \Leftrightarrow (\varphi \circ \gamma)'(0) = (\varphi \circ \eta)'(0) \tag{2.2}$$

for some chart φ of M around p. By the chain rule (2.2) holds for every chart around p and defines an equivalence relation on the set of all curves passing through p. The equivalence class $[\gamma]$ is called *(geometric) tangent vector* (of M at p). Define the *(geometric) tangent space of M at p* as the set T_pM of all geometric tangent vectors at p.

The tangents to all curves passing through a point form a vector space, the so called tangent space. It turns out to be isomorphic to the space on which M is modelled.

2.2.2 Lemma *Let φ be a d-chart of M around p, set $p_\varphi := \varphi(p)$. Then*

(a) *The map*

$$h_\varphi \colon \mathbb{R}^d \to T_pM, \quad h_\varphi(y) := [t \to \varphi^{-1}(p_\varphi + ty)]$$

is a bijection with inverse $h_\varphi^{-1} \colon T_pM \to \mathbb{R}^d, \quad [\gamma] \to (\varphi \circ \gamma)'(0)$.

(b) *For all charts φ, ψ around p, we have $h_\psi^{-1} \circ h_\varphi = \mathbf{d}(\psi \circ \varphi^{-1})(p_\varphi)$ which is an automorphism of $\mathbb{R}^d$.*

(c) *T_pM admits a unique vector space structure such that h_φ is an isomorphism for some (and hence all) charts φ of M around p.*

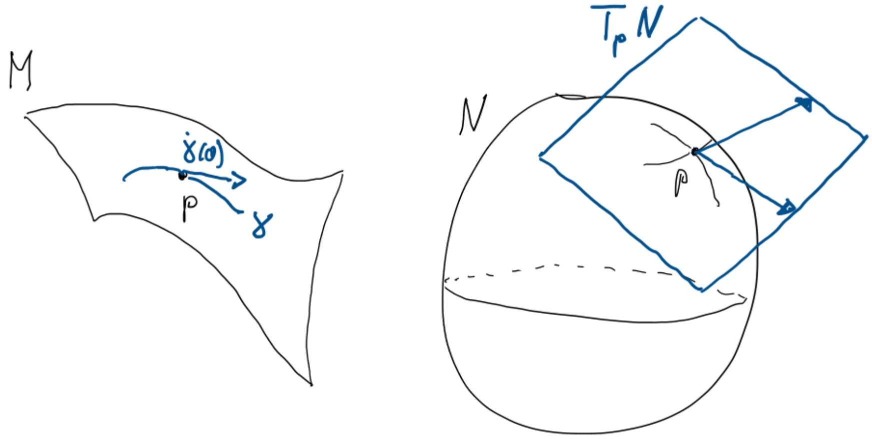


Figure 2.2.: **Left**: A point $p \in M$ and a curve γ whose derivative at p is a vector. **Right**: Two derivative vectors spanning the tangent plane T_pN at p.

Proof. (a) Note that h_φ and h_φ^{-1} are well defined and h_φ^{-1} is injective. For $y \in \mathbb{R}^d$, $h_\varphi^{-1} \circ h_\varphi(y) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \varphi(\varphi^{-1}(p_\varphi + ty)) = y$. Thus h_φ^{-1} is surjective and the inverse of h_φ.

(b)-(c) Compute for $y \in \mathbb{R}^d$: $h_\psi \circ h_\varphi^{-1}(y) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \psi(\varphi^{-1}(p_\varphi + ty)) = \mathbf{d}(\psi \circ \varphi^{-1})(p_\varphi)(y)$. Now (c) follows directly from the definition of the vector space structure. □

To study tangent spaces at different points, we will need to "straighten them out". Charts allow us to reduce to a cartesian product situation, where the tangent vectors are all mapped to a vector in the same linear space and we only need to remember the basepoint:

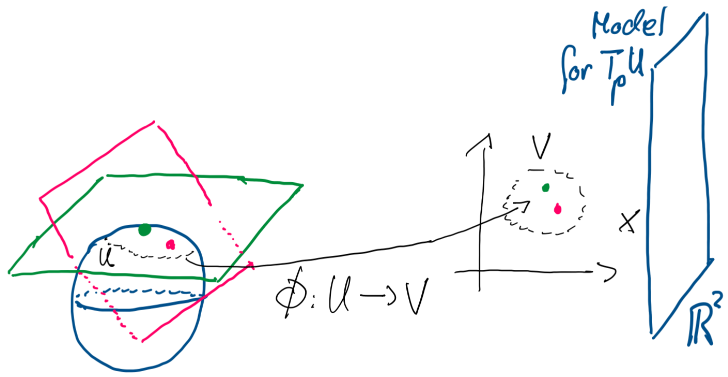


Figure 2.3.: Taking a chart φ to identify the tangent spaces at different points with one linear space (here $\mathbb{R}^2$) modelling the tangent directions.

2.2.3 Definition (tangent bundle) Let (M, A) be a d-dimensional smooth manifold. We call $TM := \coprod_{p \in M} T_pM$ the *tangent bundle* of M. Then $\pi_M : TM \to M, T_pM \ni v \mapsto p$ is called the *bundle projection*. We equip TM with the final topology with respect to the family $(T\varphi^{-1})_{\varphi \in A}$ of mappings

$$T\varphi^{-1}: V_\varphi \times \mathbb{R}^d \to TM, \quad (x, y) \mapsto [t \mapsto \varphi^{-1}(x + ty)] \in T_{\varphi^{-1}(x)}M.$$

Note that $TU_\varphi = \pi_M^{-1}(U_\varphi)$ is open in TM for all $\varphi \in A$. and

$$T\varphi := (T\varphi^{-1})^{-1}: TU_\varphi \to V_\varphi \times \mathbb{R}^d, [\gamma] \mapsto (\varphi(\gamma(0)), (\varphi \circ \gamma)^{\cdot}(0))$$

is a homeomorphism. Moreover, $B := \{T\varphi \mid \varphi \in A\}$ is a smooth atlas for TM. Thus TM becomes a smooth $2d$-dimensional manifold and $\pi_M : TM \to M$ a smooth map.

2.2.4 Example Let $U \overset{\circ}{\subseteq} \mathbb{R}^n$ be the manifold with the global chart $\iota_U : U \to \mathbb{R}^n, \iota_U(x) = x$. Taking the definitions of TU and $T\iota_U$, we get

$$T\iota_U([\gamma]) = (\gamma(0), \dot{\gamma}(0)) \in U \times \mathbb{R}^n.$$

In other words, there is a natural identification $TU = U \times \mathbb{R}^d$, so the tangent bundle here really is a cartesian product of the manifold U with $\mathbb{R}^d$.

Before we establish properties of the tangent bundle, the following definition is useful

2.2.5 Let $U \overset{\circ}{\subseteq} \mathbb{R}^d$, $d \in \mathbb{N}$ and $f : U \to \mathbb{R}^m$ a C^1-map. We define the mapping

$$Tf : U \times \mathbb{R}^d \to \mathbb{R}^m \times \mathbb{R}^m, \quad (x, v) \mapsto (f(x), \mathbf{d}f(x)(v)),$$

and call this mapping the *tangent map* of f. The chain rule becomes $T(f \circ g) = Tf \circ Tg$. The main difference with the tangent map Tf compared to $\mathbf{d}f$ is the bookkeeping. Tf remembers the base point!

2.2.6 Lemma *We check the details in Definition 2.2.3. Let $\phi, \psi \in A$:*

(a) *We have $T(\psi \circ \varphi^{-1}) \circ T\varphi = T\psi$,*

(b) *TU_ϕ is open in TM and $T\phi: TU_\phi \to V_\phi \times \mathbb{R}^d$ is a homeomorphism.*

(c) *$B = \{T\varphi \mid \varphi \in A\}$ is a smooth atlas, if M is Hausdorff, so is TM,*

(d) *π_M is a smooth map,*

Proof. (a) This follows from Lemma 2.2.2 (b): Let $[\gamma] \in TU_\varphi \cap TU_\psi$ and apply the map $T(\psi \circ \varphi^{-1}) = (\psi \circ \varphi, \mathbf{d}(\psi \circ \varphi^{-1}))$ from 2.2.5 to $T\varphi([\gamma]) = (\varphi(\gamma(0)), (\varphi \circ \gamma)^{\cdot}(0))$. Then the chain rule yields the identity. We leave the details to the reader.

(b) If $\phi, \psi \quad A$, we have $(T\psi^{-1})^{-1}(TU_\phi) = T\psi(TU_\phi \cap TU_\psi) = \psi(U_\phi \cap U_\psi) \times \mathrm{R}^d \circ V_\psi \times \mathrm{R}^d$. By the definition of the final topology TU_ϕ is open in TM.

By definition of the final topology $T\phi^{-1}$ is continuous, whence $T\phi$ is open for every $\phi \quad A$. For continuity, pick $U \circ V_\phi \times \mathrm{R}^d$ and let us show that $T(\phi)^{-1}(U)$ is open. Thus by definition of the final topology we need to show that $(T\psi^{-1})^{-1}((T\phi)^{-1}(U))$ is open for every $\psi \quad A$. Now $W := U \cap \phi(U_\phi \cap U_\psi) \times \mathrm{R}^d) \circ V_\phi \times \mathrm{R}^d$, whence a quick computation shows

$$(T\psi^{-1})^{-1}((T\phi)^{-1}(U)) = T(\psi \circ \phi^{-1})(W) \circ V_\psi \times \mathrm{R}^d$$

as $T(\psi \circ \phi^{-1})$ is a homeomorphism between open subsets of $V_\phi \times \mathrm{R}^d$ and $V_\psi \times \mathrm{R}^d$. We deduce that $(T\phi)^{-1}(U)$ is open in TM, whence $T\phi$ is continuous.

(c) By (b) each $T\varphi$ is a homeomorphism from an open subset of TM onto an open subset of $\mathrm{R}^d \times \mathrm{R}^d$, whence it is a chart. Clearly the TU_ϕ cover TM and by (a) the change of chart maps are $T(\psi \circ \phi^{-1}) = (\psi \circ \phi^{-1}, \mathbf{d}(\psi \circ \phi^{-1}))$ whence smooth. The Hausdorff property is left as Exercise 2.2.

(d) In every chart we have $\phi \circ \pi_M = \mathrm{pr}_1 \circ T\phi$, where $\mathrm{pr}_1 \colon V_\phi \times \mathrm{R}^d \to V_\phi$ is the canonical projection. As the charts conjugate π_M to a smooth map, it is smooth. □

We will now introduce tangent mappings of differentiable mappings between manifolds.

2.2.7 Definition (Tangent maps) Let $f \colon M \to N$ be a C^r-map between smooth manifolds. Then we define the mappings

$$T_p f \colon T_p M \to T_{f(p)} N, \quad [\gamma] \to [f \circ \gamma], \quad p \quad M$$

The ***tangent map*** of f is $Tf \colon TM \to TN, T_p M \quad [\gamma] \to T_p f([\gamma])$. Note that by construction $\pi_N \circ Tf = f \circ \pi_M$. Moreover, for each pair of charts ψ of N and φ of M such that $f(U_\varphi) \quad U_\psi$ the following diagram is commutative

$$\begin{array}{ccc} TU_\varphi & \xrightarrow{Tf|_{TU_\varphi}^{TU_\psi}} & TU_\psi \\ \downarrow T\varphi & & \downarrow T\psi \\ V_\varphi \times \mathrm{R}^d & \xrightarrow{T(\psi \circ f \circ \varphi^{-1})} & V_\psi \times \mathrm{R}^m \end{array}$$

This means that in the canonical charts of the tangent bundle Tf corresponds to

$$T(\psi \circ f \circ \varphi^{-1})(p, v) = (\psi \circ f \circ \varphi^{-1}(p), \mathbf{d}(\psi \circ f \circ \varphi^{-1})(p)(v)). \tag{2.3}$$

Hence the tangent map Tf is a C^{r-1}-map if f is a C^r-map. Note that the formula (2.3) identifies $T_m f$ (in the canonical charts with the derivative $\mathbf{d}(\psi \circ f \circ \varphi^{-1})(\varphi(m))$, whence $T_m f$ is a linear map by Lemma 2.2.2 (c).

2.2.8 Lemma (Chain rule on manifolds) *Let M, N, L be C^r-manifolds and $f: M \to N$, $g: N \to L$ be C^r maps. Then $T(g \circ f) = Tg \circ Tf$.*

Note that we can of course iterate the tangent construction and form the higher tangent manifolds $T^k M := \underbrace{(T(T(\cdots(T}_{k \text{ times}} M)\cdots)$ if M is a C^∞-manifold and $k \leq \infty$. Similarly one defines higher tangent maps $T^k f := \underbrace{(T(T(\cdots(T}_{k \text{ times}} f)\cdots)$.

2.2.9 Definition Let $f: M \to N$ is a smooth map and $m \in M$. The map f is called *submersive at m* if the differential $T_m f$ is surjective. Otherwise m is called a *critical point of f*.

The map f is said to be a submersion if $T_m f$ is surjective for every $m \in M$.

2.2.10 Lemma *If $f: M \to N$ is submersive at m, then there exists an open neighborhood $U \subset N$ of $p := f(m)$ and a smooth map $\sigma: U \to M$ wtih $f \circ \sigma = \mathrm{id}_U$ and $\sigma(p) = m$.*

Proof. The assertion is purely local, so without loss of generality we may assume that M is an open subset of $\mathbb{R}^d$ and N is an open subset of $\mathbb{R}^n$. Since $\tau_m: \mathbb{R}^d \to \mathbb{R}^d$, $\tau_m(x) = x - m$ is a diffeomorphims of $\mathbb{R}^d$ we can replace M with $\tau_m(M)$ and assume that $m = 0$. That f is submersive at m means that $\mathbf{d}f(m): \mathbb{R}^d \to \mathbb{R}^n$ is surjective. Let $V \subset \mathbb{R}^n$ be a linear subspace for which $\mathbf{d}f(m)|_V: V \to \mathbb{R}^m$ is a linear isomorphism. Then $F := f|_{M \cap V}: M \cap V \to N$ is a smooth map whose differential $T_m F = T_0 F$ is invertible. Hence the Inverse Function Theorem A.2.7 implies the existence of a smooth inverse σ on an open neighborhood of $p = F(m)$. □

2.2.11 Proposition (Universal property of submersions) *Let $f: M \to N$ be a surjective submersion, P a smooth manifold and $h: N \to P$ a map. Then h is smooth if and only if $h \circ f: M \to P$ is smooth. In particular, for each smooth $g: M \to P$ which is constant on all $f^{-1}(n)$, $n \in N$, there exists a unique smooth map $h: N \to P$ with $h \circ f = g$,*

Proof. If h is smooth, then also $h \circ f$ is smooth by the chain rule. Assume conversely that $h \circ f$ is smooth. To see that h is smooth, pick $n \in N$ and $n \in U \subset N$ on which there exists a smooth section $\sigma: U \to M$ with $f \circ \sigma = \mathrm{id}_U$, Lemma 2.2.10. Then $h|_U = h \circ f \circ \sigma$ is smooth on U. Since n was arbitrary, h is smooth.

The second assertion follows now immediately by defining the map $h: N \to P$ by $h(f(x)) = g(x)$, which makes sense if g is constant on the fibers of f. □

2.2.12 Corollary *If $f: M \to N$ is a bijective submersion, then f is a diffeomorphism.*

Proof. Apply Proposition 2.2.11 to $P := M$ and $g = \mathrm{id}_M$, then $h = f^{-1}$. □

2.2.13 Remark The smooth map $f : M := \mathbb{R} \to N =: \mathbb{R}, f(x) = x^3$ is submersive at all points $x = 0$. $g = \mathrm{id}_{\mathbb{R}} : M \to \mathbb{R} =: P$ is smooth, bijective and constant on the fibers of f, but the map $\hat{g} : N \to P, \hat{g}(x) = \sqrt[3]{x}$ is not smooth at 0. Hence the assumption of Proposition 2.2.11 that f is a submersion is really needed.

2.2 Exercises

1. Verify the details in Definition 2.2.7 and Lemma 2.2.8. Show in particular that

a) Tf is well defined and defines a C^{r-1}-map with the claimed properties.

b) Check that for $M = U \circ \mathbb{R}^d$, both definitions of tangent mappings coincide.

c) the manifold TM is a Hausdorff topological space.

Hint: Consider two cases for $v, w \ \ TM$: $\pi_M(v) = \pi_M(w)$ and $\pi_M(v) = \pi_M(w)$.

2. Show that for smooth manifolds $M_1, \ldots, M_n$ the projection maps

$$\pi_i : M_1 \times \cdots \times M_n \to M_i, \quad (x_1, \ldots, x_n) \to x_i \text{ induce a diffeomorphism}$$
$$(T\pi_1, T\pi_2, \ldots, T\pi_n) : T(M_1 \times \cdots \times M_n) \to TM_1 \times TM_2 \times \cdots TM_n.$$

3. Let $\mu : E \times F \to W$ be a bilinear map and M a smooth manifold. For $f \ \ C^\infty(M, E), g \ \ C^\infty(M, F)$ and $p \ \ M$ set $h(p) := \mu(f(p), g(p))$. Show that h is smooth with $T(h)v = \mu(T(f)v, g(p)) + \mu(f(p), T(g)v)$ for $v \ \ T_p(M)$.

4. Inverse function theorem for manifolds Let $f : M \to N$ be smooth and $p \ \ M$ such that $T_pf : T_pM \to T_pN$ is a linear isomorphism. Show that there exists $p \ \ U \circ M$ such that $f|_U : U \to f(U)$ is a diffeomorphism onto an open subset of N.

2.3. Vector fields, their Lie bracket and differential equations

In this section we study special smooth maps from a manifold into its tangent bundle, the so called vector fields. As in the vector space case, vector fields are the right type of mapping to define differential equations on manifolds. They will turn out to be a useful tool. We only treat smooth vector fields as only these vector fields yield a Lie algebra.

2.3.1 Definition Let M be a smooth manifold. A smooth map $X : M \to TM$ is called *vector field* if $\pi_M \circ X = \mathrm{id}_M$. We write $V(M)$ for the set of all smooth vector fields.

The next lemma is a consequence of the fact that for every $p \in M$ a vector field takes its image in T_pM and we know that the tangent space is a vector space.

2.3.2 Lemma *Let M be a smooth manifold, then $\mathcal{V}(M)$ is a vector space with respect to the pointwise operations*

$$(X+Y)(p) := X(p) + Y(p), \quad (cX)(p) = c(X(p)) \text{ in } T_pM.$$

More generally, we can multiply vector fields with smooth functions:

$$(fX)(p) := f(p)X(p), \quad f \in C^\infty(M, \mathbb{R}), X \in \mathcal{V}(M).$$

The following example shows how the concept of vector fields on a manifold is related to the concept of a vector field on an open subset of a vector space (which we discussed in Appendix A.3)

2.3.3 Example If $U \overset{\circ}{\subseteq} \mathbb{R}^d$, we have $TU = U \times \mathbb{R}^d$ and $\pi_U : U \times \mathbb{R}^d \to U, (u,e) \mapsto u$, by Example 2.2.4. Thus a vector field $X \in \mathcal{V}(U)$ is given as $X = (X_U, X_2) \colon U \to U \times \mathbb{R}^d$ and we must have $X_U = \mathrm{id}_U$. Hence a vector field on U is uniquely determined by the smooth map $X_2 \in C^\infty(U, \mathbb{R}^d)$. We call X_2 the *principal part* of X. For any $U \overset{\circ}{\subseteq} \mathbb{R}^d$, the map

$$\mathcal{V}(U) \to C^\infty(U, \mathbb{R}^d), \quad X = (\mathrm{id}_U, \tilde{X}) \mapsto \tilde{X}$$

is a linear isomorphism. The difference between the two concepts of vector fields is bookkeeping, on manifolds we need to remember the base point.

2.3.4 (Principal part in a chart) Let M be a manifold and $X \in \mathcal{V}(M)$. Then we have an analogue of the principal part for every d-chart (φ, U_φ) of M. Localising, we obtain

$$T\varphi \circ X \circ \varphi^{-1} = (\mathrm{id}_{V_\varphi}, X_\varphi) \text{ for } X_\varphi := \mathrm{pr}_2 \circ T\varphi \circ X \circ \varphi^{-1} \colon V_\varphi \to \mathbb{R}^d$$

where pr_2 the projection onto the second component. We call the smooth map X_φ the *local representative of X* or *the principal part of X* with respect to φ.

We will use the notion of integral curves and flows for vector fields, whence we recall the definition of these objects.

2.3.5 For a C^1-curve $c \colon (a,b) \to M$ we can define $\dot{c}(t) := T_tc(1) \in T_{c(t)}M$, where we use $T(a,b) = (a,b) \times \mathbb{R}$. Let $X \in \mathcal{V}(M)$, we say a C^1-curve $c \colon (a,b) \to M$ is an *integral curve* for X if for every $t \in (a,b)$ the curve satisfies $\dot{c}(t) = X(c(t))$.

It follows from the theory of ordinary differential equations that for every $m \in M$ there exists an integral curve c_m of X on some open interval $J_m :=]-\varepsilon, \varepsilon[$ such that $c_m(0) = m$. Indeed we can just localize the equation and the vector field using charts and then use the theory of ODEs recalled in Appendix A.3. Moreover, the *flow*

$$\mathrm{Fl}^X : \bigcup_{m \in M} \{m\} \times J_m \to M, (m,t) \mapsto c_m(t),$$

defines a continuous map on some open subset of $M \times \mathbb{R}$. If Fl^X is defined on all of $M \times \mathbb{R}$, then we call X *complete*.

We just state the following meta theorem which is based on the observation that we can always localise vector fields and their flows in charts. There the local solution theory together with uniqueness of the solution allows us to patch together the local solutions and obtain solutions on the manifold. For details and proofs we refer to [HN12].

2.3.6 Theorem *Let M be a manifold. Any vector field $X \in \mathcal{V}(M)$ admits a unique smooth flow $\mathrm{Fl}^X : D_X \to M$, where D_X is an open subset of $M \times \mathbb{R}$. Moreover, if X smoothly depends on parameters (cf. Proposition A.3.13), then so does the flow Fl^X.*

Relatedness and the Lie algebra of vector fields

In the rest of the section we will construct the Lie algebra of vector fields on a smooth manifold. To introduce a Lie bracket on vector fields it is useful to first consider the notion of relatedness of vector fields. This will allow us to transport constructions in charts to the manifold.

2.3.7 Definition Let $f : M \to N$ be smooth. We call the vector fields $X \in \mathcal{V}(M), Y \in \mathcal{V}(N)$ *f-related* if $Y \circ f = Tf \circ X$.

Note that in general there is no reason why for a given map f and a vector field X that there is an f related vector field Y, cf. Exercise 2.3 1.

2.3.8 Lemma *Let M be a d-dimensional manifold with atlas $\mathcal{A}$. Let $(X_\varphi)_{(\varphi,U_\varphi) \in \mathcal{A}}$ be a family of smooth maps $X_\varphi \colon \varphi(U_\varphi) \to \mathbb{R}^d$ such that every pair of vector fields $(\mathrm{id}_{\varphi(U_\varphi)}, X_\varphi), (\mathrm{id}_{\psi(U_\psi)}, X_\psi)$ is $\psi \circ \varphi^{-1}$-related on $\varphi(U_\psi \cap U_\varphi)$. Then there is a unique vector field $X \in \mathcal{V}(M)$ whose principal part with respect ot φ coincides with the X_φ.*

Proof. Define $X\colon M \to TM, p \to T\varphi^{-1}(\varphi(p), X_\varphi(\varphi(p)))$ for $p \in U_\varphi$. Since the maps X_φ, X_ψ are related by the change of charts on the overlap $U_\varphi \cap U_\psi$, the mapping is well defined. By construction it is smooth and a vector field. □

2.3.9 For a vector field $X \in V(M)$ and $f \in C^\infty(M, \mathbb{R}^m), m \in \mathbb{N}$ we define the *Lie derivative* $L_X(f) \in C^\infty(M, \mathbb{R}^m)$ via

$$L_X(f)(m) \coloneqq \mathrm{pr}_2 \circ Tf \circ X(m). \tag{2.4}$$

Further, if now $X, Y \in V(U)$, $U \subseteq \mathbb{R}^d$ with principal parts $\tilde{X}, \tilde{Y}$ and define

$$[\tilde{X}, \tilde{Y}] \coloneqq L_X(\tilde{Y}) - L_Y(\tilde{X}), \quad [\tilde{X}, \tilde{Y}](p) = \mathbf{d}\tilde{Y}(p)(\tilde{X}(p)) - \mathbf{d}\tilde{X}(p)(\tilde{Y}(p)), p \in U. \tag{2.5}$$

Note that the bracket Equation (2.5) makes sense on $C^\infty(U, \mathbb{R}^d)$ and we will see that gives rise to a Lie bracket on $C^\infty(U, \mathbb{R}^d)$. Our strategy will then be to transport this structure to $V(M)$. For this we need the following:

2.3.10 Lemma *Let $U \subseteq \mathbb{R}^d$, $V \subseteq \mathbb{R}^n$ be open and $f \in C^\infty(U, V)$, $X_1, X_2 \in C^\infty(U, \mathbb{R}^d)$ and $Y_1, Y_2 \in C^\infty(V, \mathbb{R}^n)$. Assume that X_i is f-related to Y_i for $i = 1, 2$, then $[X_1, X_2]$ is f-related to $[Y_1, Y_2]$.*

Proof. To compute $\mathbf{d}f(p)([X_1, X_2](p))$, note that by relatedness $Y_i(f(p)) = \mathbf{d}f(p)(X_i(p))$. Compute the bracket with (2.5) and use the chain rule with (A.3), then insert the identity. For $i = 1, 2$, $(x, v) \in U \times \mathbb{R}^d$ this yields

$$\mathbf{d}f(p)(\mathbf{d}X_i(p)(v)) = \mathbf{d}Y_i(f(p))(\mathbf{d}f(p)(v)) - \mathbf{d}^2 f(p)(X_i(p), v) \tag{2.6}$$

Note that by inserting for v one of the vectors $X_i(p)$ we can use (2.6) to obtain from the right hand side of (2.5) the following

$$\begin{aligned}
\mathbf{d}f(p)([X_1, X_2](p)) &= \mathbf{d}f(p)(\mathbf{d}X_2(p)(X_1(p))) - \mathbf{d}f(p)(\mathbf{d}X_1(p)(X_2(p))) \\
&= \mathbf{d}Y_2(f(p))(\mathbf{d}f(p)(X_1(p))) - \mathbf{d}^2 f(p)(X_2(p), X_1(p)) \\
&\quad - \mathbf{d}Y_1(f(p))\,\mathbf{d}f(p)(X_2(p))) + \mathbf{d}^2 f(p)(X_1(p), X_2(p)) \\
&= \mathbf{d}Y_2(f(p))(Y_1(f(p))) - \mathbf{d}Y_1(f(p))(Y_2(f(p))) = [Y_1, Y_2](f(p))
\end{aligned}$$

where the second order derivatives cancel by Schwarz' Theorem A.2.4. □

2.3.11 Proposition *Let M be a smooth manifold. Then for every $X, Y \in V(M)$ there exists a vector field $[X, Y] \in V(M)$ which is uniquely determined by the property that for each chart (ϕ, U) of M the following equation holds for the principal parts*

$$[X, Y]_\phi = [X_\phi, Y_\phi].$$

Proof. If $(\phi, U), (\psi, V)$ are charts of M, the vector fields $(\mathrm{id}_{\phi(U)}, X_\phi)$ on $\phi(U)$ and $(\mathrm{id}_{\psi(V)}, X_\psi)$ on $\psi(V)$ are $(\psi \circ \phi^{-1})$-related. Therefore, Lemma 2.3.10 implies that $[X_\phi, Y_\phi]$ is $(\psi \circ \phi^{-1})$-related to $[X_\psi, Y_\psi]$. This in turn is equivalent to

$$T\phi^{-1} \circ (\mathrm{id}_{\phi(U)}, [X_\phi, Y_\phi]) \circ \phi = T\psi^{-1} \circ (\mathrm{id}_{\psi(V)}, [X_\psi, Y_\psi]) \circ \psi$$

on the open subset $U \cap V$. Hence there exists a unique vector field $[X, Y] \in V(M)$ satisfying

$$[X, Y]|_U = T\phi^{-1} \circ (\mathrm{id}_{\phi(U)}, [X_\phi, Y_\phi]) \circ \phi$$

for each chart (ϕ, U), i.e. $[X, Y]_\phi = [X_\phi, Y_\phi]$. □

The Lie algebra property for vector fields with the Lie bracket will be established using a general definition we give now.

2.3.12 Definition Let $(A, \cdot)$ be an associative algebra,[2] then by Exercise 1.5 3. $\mathrm{Lin}(A, A)$ forms a Lie algebra under the commutator bracket $[\varphi, \psi] := \varphi \circ \psi - \psi \circ \varphi$.

A mapping $\varphi \in \mathrm{Lin}(A, A)$ is called *derivation* of the algebra A if it satisfies the Leibniz rule

$$\varphi(a \cdot b) = \varphi(a) \cdot b + a \cdot \varphi(b) \quad \forall a, b \in A.$$

We denote by $\mathrm{der}(A)$ the *set of all derivations* of A and note that it forms a Lie subalgebra of $(\mathrm{Lin}(A, A), [\cdot, \cdot])$ (Exercise!).

To explain the notation consider $C^\infty(U, \mathbb{R})$ for $U \overset{\circ}{\subset} \mathbb{R}^d$. The pointwise multiplication turns $C^\infty(U, \mathbb{R})$ into an associative algebra. Then L_X for $X \in V(U)$ is linear in f and satisfies the Leibniz rule of a derivation:

$$L_X(f \cdot g) = L_X(f) \cdot g + f \cdot L_X(g). \tag{2.7}$$

With other words, L_X is a derivation of the algebra $C^\infty(U, \mathbb{R})$.

2.3.13 Lemma *Let $U \overset{\circ}{\subset} \mathbb{R}^d$ and $X, Y \in C^\infty(U, \mathbb{R}^d) = V(U)$. We abuse notation and will write the Lie derivative of the vector field (id_U, X) simply as L_X. Then*

(a) $L_{[X,Y]} = L_X \circ L_Y - L_Y \circ L_X = [L_X, L_Y]$

(b) *The map $L\colon C^\infty(U, \mathbb{R}^d) \to \mathrm{der}(C^\infty(U, \mathbb{R})), X \mapsto L_X$ is linear and injective*

(c) *The map $[\cdot, \cdot]\colon C^\infty(U, \mathbb{R}^d) \times C^\infty(U, \mathbb{R}^d) \to C^\infty(U, \mathbb{R}^d), [X, Y] = L_X(Y) - L_Y(X)$ turns the space $C^\infty(U, \mathbb{R}^d)$ into a Lie algebra.*

[2] An associative $\mathbb{K}$-algebra $(A, \cdot)$ is a $\mathbb{K}$-vector space together with a bilinear map $\cdot\colon A \times A \to A$, the product, which is associative, i.e. $(a \cdot b) \cdot c = a \cdot (b \cdot c), \ \forall a, b, c \in A$.

Proof. (a) From (A.3) we deduce

$$L_X(L_Y(f))(p) = \mathbf{d}^2 f(p)(Y(p), X(p)) + \mathbf{d}f(p)(\mathbf{d}Y(p)(X(p))), \quad p \in U.$$

Interchanging X and Y, the second order terms cancel by Schwarz' Theorem A.2.4

$$[L_X, L_Y](f)(p) = L_{[X,Y]}(f)(p).$$

(b) L_X is linear in X as $\mathbf{d}f(x)$ is. Thus it suffices to prove that the kernel of L is trivial. Let $X \in C^\infty(U, \mathbb{R}^d)$ be a map with $X(x) \neq 0$ for some $x \in U$. Taking $\lambda_I : \mathbb{R}^d \to \mathbb{R}$ as the projection onto the ith componentwe can choose i such that $\lambda_i(X(x)) \neq 0$. Then $L_X(\lambda_i)(x) = \mathbf{d}\lambda_i(x)(X(x)) = \lambda_i(X(x)) \neq 0$ and thus $L_X \neq 0$.

(c) We have to check Definition 1.5.6: Clearly $[\cdot,\cdot]$ is bilinear, and $[X, X] = 0$ holds, so (L1) is satisfied. For (L2), we write the Jacobi identity in the equivalent form of a sum of cyclically permuted Lie brackets. We have thus to check that $\sum_{\text{cycl}}[X,[Y,Z]]$ vanishes for all $X, Y, Z \in C^\infty(U, \mathbb{R}^d)$. However,

$$L\left(\sum_{\text{cycl}}[X,[Y,Z]]\right) = \sum_{\text{cycl}}[L_X,[L_Y,L_Z]] = 0$$

where we have used linearity of L from (b) and that by (a), L forms a Lie algebra morphism as well as that $\operatorname{der}(C^\infty(U,\mathbb{R}))$ forms a Lie algebra. Since L is injective by (b), we see that the Jacobi identity holds. □

Note that the identities (L1) and (L2) of the Lie bracket $[X, Y]$ of vector fields $X, Y \in \mathcal{V}(M)$ can be checked locally on a chart domain. Hence, the following is a consequence of Lemma 2.3.13, we leave it as an exercise.

2.3.14 Proposition *$(\mathcal{V}(M), [\cdot,\cdot])$ is a Lie algebra.*

Again the Lie bracket of vector fields measures how much vector fields "commute with each other". This is a condition on the flow of the associated vector fields.

2.3.15 Lemma *If $X, Y \in \mathcal{V}(M)$ are complete vector fields then*

$$\mathrm{Fl}_t^X \circ \mathrm{Fl}_s^Y = \mathrm{Fl}_s^Y \circ \mathrm{Fl}_t^X, \quad \forall s,t \in \mathbb{R} \text{ if and only if } [X,Y] = 0.$$

We shall not provide a proof of Lemma 2.3.15 as it needs a few more technical steps. One can prove that for a vector field X and a smooth function $f \in C^\infty(M,\mathbb{R})$, the identity

$$L_X(f) = \lim_{t\to 0}(f \circ \mathrm{Fl}_t^X - f)$$

holds. Expanding this formula suitably one can write the Lie bracket as a Lie derivative and this leads to a proof of Lemma 2.3.15 as in [HN12, Corollary 8.5.17].

2.3 Exercises

1. We investigate the notion of relatedness of vector fields.

a) Let $f: M \to N$ be a diffeomorphism. Show that for every vector field $X \in V(M)$ there is an f-related vector field on N.

b) Let $M = (-1, 0) \cup (0, 1)$ and $N = (0, 1)$ with the usual manifold structure as open sets of $\mathbb{R}$. Show that for $g: M \to N, g(x) := \begin{cases} x+1, & x \in (-1,0) \\ x, & x \in (0,1) \end{cases}$ there are vector fields on M which do not admit g-related vector fields on N.

2. Let X, Y be vector fields on $\mathbb{R}^n$. We work exclusively with principal parts $\tilde{X}, \tilde{Y}$.

a) Show that the Lie bracket from Equation (2.5) satisfies

$$\widehat{[X, Y]}(z) = J_{\tilde{X}}(z) \cdot \tilde{Y}(z) - J_{\tilde{Y}}(z) \cdot \tilde{X}(z),$$

where $J_{\tilde{X}}$ is the Jacobian-matrix associated to the principal part.

b) Determine all pairs of smooth functions $f, g: \mathbb{R}^2 \to \mathbb{R}$ such that $[V, W] = 0$ for the vector fields $V(x, y) := \begin{pmatrix} f(x,y) \\ g(x,y) \end{pmatrix}$ and $W(x, y) := \begin{pmatrix} e^y \\ 0 \end{pmatrix}$ on $\mathbb{R}^2$.

3. Show that for an associative algebra $(A, \cdot)$, $\mathrm{der}(A)$ forms a Lie subalgebra of $(\mathrm{Lin}(A, A), [\cdot, \cdot])$.

4. Use Lemma 2.3.13 to prove Proposition 2.3.14.
Hint: Can we localise (L1) and (L2) from Definition 1.5.6 in charts?

5. We investigate the Lie bracket and its associated Lie derivative. Let M, N be smooth manifolds. Let $\phi: N \to M$ be a smooth map. We define the pullback map

$$\phi^* : C^\infty(M, \mathbb{R}) \to C^\infty(N, \mathbb{R}), \quad f \mapsto f \circ \phi.$$

a) Show that if X and Y are ϕ-related, then $L_Y \circ \phi^* = \phi^* \circ L_X$.

b) For $f \in C^\infty(M, \mathbb{R})$ use a) to show $L_{[X,Y]}(f) = L_X(L_Y(f)) - L_Y(L_X(f))$.
Note: This identity also follows from Lemma 2.3.13. Why?

6. **(Related Vector Field Lemma, global form)** Let $\phi: M \to N$ be a smooth map and $X, X' \in V(M)$ Show that if Y and Y' are ϕ-related to X and X', then $[X', Y']$ and $[X, Y]$ are ϕ-related.
Hint: This is a local question, so use charts to go back to Lemma 2.3.10.

3. Lie groups beyond matrix groups

In this chapter we define and study the general notion of a Lie group. This builds on the previous sections on manifolds and defines a Lie group as a group which also carries a smooth manifold structure. Here is the definition.

3.0.1 Definition A *Lie group* G is a manifold G endowed with a group structure such that the multiplication map $m_G \colon G \times G \to G$ and the inversion map $\iota \colon G \to G$ are smooth. A morphism of Lie groups is a smooth group homomorphism.

Standard notation

Let us fix some standard notation for objects occurring frequently in conjunction with Lie groups. Let G be a Lie group, we shall write

- $\mathbf{1}_G$ for the unit element (or shorter $\mathbf{1}$),
- m_G for multiplication, ι_G for inversion,
- for $g \in G$ we let $\lambda_g \colon G \to G, h \mapsto gh$ and $\rho_g \colon G \to G, h \mapsto hg$ the *left (right) translation*. (Observe that $\lambda_g(\rho_h(x)) = gxh = \rho_h(\lambda_g(x))$.)

3.0.2 Example The vector space $\mathbb{R}^d$ is a Lie group with respect to vector addition and the usual manifold structure.

3.0.3 Example It follows directly from Proposition 1.1.2 that for every $n \in \mathbb{N}$ and $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ the group $\mathrm{GL}_n(\mathbb{K})$ is a Lie group.

3.0.4 Example (Circle and torus group) (a) We already know that the circle

$$\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

is a submanifold of $\mathbb{R}^2$ (cf. Example 2.1.5). Identifying it with $\{z \in \mathbb{C} \mid |z| = 1\}$, it also inherits a group structure, given by

$$(x, y) \cdot (x', y') := (xx' - yy', xy' + x'y) \text{ and } (x, y)^{-1} = (x, -y).$$

You will verify in Exercise 3.1 1. that $\mathbb{S}^1$ is a Lie group.

(b) (n-dimensional torus) In view of (a) we have a natural manifold structure on the *n-dimensional torus* $\mathbb{T}^n := (\mathbb{S}^1)^n$. The corresponding direct product group structure

$$(t_1, \dots, t_n) \cdot (s_1, \dots, s_n) := (t_1 s_1, \dots, t_n s_n)$$

turns $\mathbb{T}^n$ into a Lie group Exercise 3.1 2.

3.0.5 Definition Let G be a Lie group. We call a submanifold $H \subseteq G$ a *Lie subgroup* if it is a subgroup of G.

Let us clarify that Lie subgroups inherit a Lie group structure from the group they sit in.

3.0.6 Lemma *If G is a Lie group and H a Lie subgroup, then H is also a Lie group in the sense of Definition 3.0.1.*

Proof. Let $H \subseteq G$ be a Lie subgroup. Then by definition H is a subgroup and a submanifold, so H inherits a manifold structure. We need to prove that the group operations $m_H : H \times H \to H$ and $\iota_H : H \to H$ are smooth with respect to this manifold structure. If $i_H : H \to G$ is the inclusion, then Lemma 2.1.16 shows that i_H is smooth. Moreover,

$$i_H \circ m_H = m_G \circ (i_h \times i_H) \text{ and } i_H \circ \iota_H = \iota_G \circ i_H.$$

Now since G is a Lie group, the maps on the right hand side are smooth and factor through the submanifold H, whence by Lemma 2.1.16 m_H and ι_H are smooth. We deduce that H is a Lie group in the sense of Definition 3.0.1. □

3.0.7 Proposition *Every linear Lie group G is a Lie group in the sense of Definition 3.0.1. Moreover, every linear Lie group is a Lie subgroup of $\mathrm{GL}_n(\mathbb{K})$ (for the choice of $\mathbb{K}$ for which G is realised as closed subgroup of $\mathrm{GL}_n(\mathbb{K})$*

Proof. By Lemma 3.0.6 all we have to do is to prove that G is a Lie subgroup of $\mathrm{GL}_n(\mathbb{K})$, i.e. we need to show that G is a submanifold of $\mathrm{GL}_n(\mathbb{K})$.

Step 1: A chart around the identity By Proposition 1.4.9 there are open neighborhoods $0 \in U \subseteq^\circ M_n(\mathbb{K})$ and $I_n \in V \subseteq^\circ \mathrm{GL}_n(\mathbb{K})$ such that the matrix exponential is a diffeomorphism $\exp : U \to V$. Its inverse $\log : V \to U$ is thus a manifold chart for $\mathrm{GL}_n(\mathbb{K})$. Now G is a linear subgroup, whence Proposition 1.7.3 shows that after possibly shrinking V we have $\log(V \cap G) = \mathbf{L}(G) \cap U$. Since $\mathbf{L}(G)$ is a real subspace of $M_n(\mathbb{K})$ this shows that $\phi := \log|_{V \cap G}^{U \cap \mathbf{L}(G)}$ is a submanifold chart for G around the identity.

The idea is now to use the multiplication to transport the chart around the identity to other points. This will yield submanifold charts at any other point in G.

Step 2: Submanifold charts for every $g \in G$. Let $g \in G$. Then we define

$$\Phi_g \colon gV = \{X \in \mathrm{GL}_n(\mathbb{K}) \mid X = gv, v \in V\} \to U, \Phi_g(x) = \phi(g^{-1}x).$$

Since multiplication in $\mathrm{GL}_n(\mathbb{K})$ is smooth, Φ_g is a smooth diffeomorphism. By construction it restricts to a submanifold chart with $\Phi_g(gV \cap G) = U \cap \mathbf{L}(G)$. These submanifold charts form an atlas of submanifold charts for G (Exercise to check the details). □

Now that we know that all linear Lie groups are Lie groups in the more general sense, the question arises: Are all Lie groups linear Lie groups? The next example gives a negative answer:

3.0.8 Example (The Heisenberg quotient) Consider the set $Q := \mathbb{R}^2 \times \mathbb{S}$. Since $\mathbb{R}^2$ is a vector space, it is a manifold and $\mathbb{S}$ is a manifold (even a 1-dimensional submanifold of $\mathbb{R}^2$) by Example 2.1.5. Hence Q is a 3-dimensional submanifold of $\mathbb{R}^2 \times \mathbb{R}^2$ (cf. Definition 2.1.13). We define a multiplication on Q via

$$\mu \colon Q \times Q \to Q, \quad \mu((x,y,z),(a,b,c)) := (x+a, y+b, z\cdot c \cdot e^{ixb}) \tag{3.1}$$

where "·" is the product from Example 3.0.4 (a). One can show that Q with this multiplication is a non-commutative Lie group whose inverse is given by the formula

$$(x,y,z)^{-1} = \left(-x, -y, \frac{1}{z}e^{ixy}\right).$$

We call Q the *Heisenberg quotient* since it can be obtained as the quotient of the Heisenberg group modulo the cyclic subgroup generated by the matrix $\begin{pmatrix} 1 & 0 & 2\pi \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The Heisenberg quotient is a 3-dimensional Lie group and the only example we have seen so far which is *NOT* isomorphic to a subgroup of $\mathrm{GL}_n(\mathbb{K})$ for some $n \in \mathbb{N}$. The proof for this statement uses some arguments from linear algebra and is contained in Appendix C.

For Lie groups, the tangent bundle TG is again a Lie group and moreover, the tangent bundle is trivial (i.e. it splits as a product product manifold where one of the manifolds is a vector space and the other the base manifold G).

3.0.9 Lemma *Let G be a Lie group. Identify $T(G \times G) = TG \times TG$.*

(a) *The tangent map of the multiplication*

$$\begin{aligned} Tm_G \colon T(G\times G) = TG\times TG \to TG, \\ T_gG\times T_hG \ni (v_g, w_h) \mapsto Tm_G(v_g,w_h) = T_g\rho_h(v_g) + T_h\lambda_g(w_h) \end{aligned} \tag{3.2}$$

induces a Lie group structure on TG with identity element $0_{\mathbf{1}} \in T_{\mathbf{1}}G$ and inversion

$$T\iota_G \colon TG \to TG,\ T_gG \ni v \mapsto -T\rho_{g^{-1}}T\lambda_{g^{-1}}(v) = -T\lambda_{g^{-1}}T\rho_{g^{-1}}(v). \tag{3.3}$$

The projection $\pi_G \colon TG \to G$ becomes a morphism of Lie groups with kernel $(T_{\mathbf{1}}G, +)$ and the zero-section $\mathbf{0}\colon G \to TG, g \mapsto 0_g$ is a morphism of Lie groups with $\pi_G \circ \mathbf{0} = \mathrm{id}_G$. In particular, (3.2) and (3.3) yield for $x, y \in T_{\mathbf{1}}G$ the formulae

$$T_{(\mathbf{1},\mathbf{1})}m_G(x,y) = x + y \quad T_{\mathbf{1}}\iota_G(x) = -x.$$

(b) *The map*

$$\Phi \colon G \times T_{\mathbf{1}}G \to TG, \quad (g, v) \mapsto g.v := Tm_G(0_g, v)$$

is a diffeomorphism.

Proof. (a) Since m_G and ι_G are smooth, the same holds for their tangent maps. The group axioms for TG follow from the ones for G by virtue of the chain rule (and has the claimed unit element). Linearity of the tangent map implies (3.2), we leave this formula and (3.3) as Exercise 3.1 3.. From the definition of the zero section and the projection, the morphism properties follow. As a result of (3.2), $T_{(\mathbf{1},\mathbf{1})}m_G(v_{\mathbf{1}}, w_{\mathbf{1}}) = v_{\mathbf{1}} + w_{\mathbf{1}}$. Hence the multiplication on the normal subgroup $\ker \pi_G = T_{\mathbf{1}}G$ is the addition.

(b) Since $\Phi = Tm_G(0_g, v) = Tm_G(\mathbf{0}(g), v)$ and the zero-section $\mathbf{0}$ is smooth, smoothness of the multiplication shows that Φ is smooth. Now a computation shows that $\Phi^{-1}(v) = (\pi_G(v), T_{\pi_G(v)}\lambda_{\pi_G(v)^{-1}}(v))$, whence Φ is bijective and its inverse is smooth (as inversion and multiplication in G are smooth and the projection π_G is smooth). □

3.0.10 Remark Note that Lemma 3.0.9 shows that $T_{\mathbf{1}}G$ is a normal Lie subgroup of TG. Moreover, instead of left multiplication one can as well use right multiplication to identify the tangent bundle (the two different choices are related by the adjoint action (see Definition 3.1.9 below).

The tangent space at the identity of a Lie group plays a special rôle. In the next section this tangent space will be endowed with an additional structure, the Lie bracket turning it into a Lie algebra. We remark that contrary to the treatment of linear Lie groups in section 1, we here define the Lie algebra first. Afterwards we will treat the Lie group exponential.

3.1 Exercises

1. Show that the natural group structure on $\mathbb{T} = \mathbb{S}^1 \subset \mathbb{R}^2$ turns the circle group from Example 3.0.4 into a Lie group.

2. Let $\phi\colon G \to H$ be a group morphism, G, H Lie groups. Show that ϕ is a Lie group morphism if and only if there is a $\mathbf{1}$-neighborhood U such that $\phi|_U$ is smooth.

3. Let $G_1, \ldots, G_n$ be Lie groups and $G := G_1 \times G_2 \times \cdots \times G_n$ be endowed with the product manifold structure and the direct product group structure. Show that G is a Lie group.

4. Check the missing details in the proof of Proposition 3.0.7 (notation as in the proof):

a) Argue that $gV = \{X \in \mathrm{GL}_n(\mathbb{K}) \mid X = gv, v \in V\}$ is open if V is open.

b) Compute Φ_g^{-1} and argue why Φ_g is a diffeomorphism which restricts to a diffeomorphism $gV \cap G \to U \cap \mathbf{L}(G)$ if $g \in G$.

c) Compute the change of charts $\Phi_g \circ \Phi_h^{-1}$ and prove that it is smooth.

5. Work out the missing details in the proof of Lemma 3.0.9.

a) Prove (3.2) and verify that the tangent maps induce a group structure on TG.

b) Establish (3.3), i.e. $T_a\iota(v) = -T\lambda_{a^{-1}}T\rho_{a^{-1}}(v) = -T\rho_{a^{-1}}T\lambda_{a^{-1}}(v)$.
Hint: Let $\gamma\colon]-\varepsilon, \varepsilon[\to G$ be smooth with $\gamma(0) = a$. Differentiate the relation $\mathbf{1} = \gamma(t)(\gamma(t))^{-1}$.

c) Show that one can obtain a diffeomorphism $TG = T_{\mathbf{1}}G \times G$ using right multiplication instead of left multiplication in part (b).

6. We revisit the Heisenberg quotient Q from Example 3.0.8.

a) Prove that Q is indeed a Lie group with the multiplication and inverse given in the example. **Hint:** By Exercise 3.1.1 we know that $\mathbb{S}^1$ is a Lie group.

b) Let H be the Heisenberg group from Example 1.1.13. Show that the map

$$q\colon H \to Q, \quad \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \mapsto (x, y, e^{iz})$$

is a smooth group morphism.
Hint: By Proposition 3.0.7 it suffices that q is a group morphism and that $q \circ \exp$ is smooth (exp is the matrix exponential), then use Exercise 1.9 1. b).

c) Show q is surjective and its kernel is the cyclic subgroup D generated by $\begin{pmatrix} 1 & 0 & 2\pi \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Then deduce that indeed $H/D = Q$ as groups.

3.1. The Lie algebra of a Lie group

We associate now to an abstract Lie group a Lie algebra. This construction allows one to reformulate many problems in Lie theory in terms of linear algebra. Contrary to the treatment for matrix Lie groups we can not start by using the matrix exponential function (as we do not have a suitable replacement for an abstract Lie group).

3.1.1 Let G be a Lie group. A vector field $X \quad V(G)$ is called *(left) invariant* if X is λ_g-related to itself for all $g \quad G$ (i.e. $X \circ \lambda_g = T\lambda_g \circ X$). We write $V\ (G)$ for the set of left-invariant vector fields. $V\ (G)$ is a Lie subalgebra of $V(G)$.

3.1.2 Proposition *Let G be a Lie group, then the map*

$$\Theta \colon T_1 G \to V\ (G), \quad v \to (g \to T_1\lambda_g(v))$$

is a linear isomorphism with inverse $\Theta^{-1}(X) = X(\mathbf{1})$. Thus $\mathbf{L}(G) := T_1 G$ can be endowed with the Lie bracket $[v, w] := \Theta^{-1}([\Theta(v), \Theta(w)]) = [\Theta(v), \Theta(w)](\mathbf{1})$ turning it into a Lie algebra. We call $(\mathbf{L}(G), [\cdot, \cdot])$ the Lie algebra associated to *G.*

Proof. As $\Theta(v)(hg) = T_1\lambda_{hg}(v) = T_g\lambda_h T_1\lambda_g(v) = T_g\lambda_h\Theta(v)(g)$, the map Θ makes sense and its image consists of left-invariant vector fields. Linearity of Θ follows directly from the linearity of the tangent map. For $X \quad V\ (G)$ we have $X(g) = X \circ \lambda_g(\mathbf{1}) = T_1\lambda_g X(\mathbf{1}) = \Theta(X(\mathbf{1}))(g)$, Θ is surjective. As the translations λ_g are diffeomorphisms, it is clear that only $0 \quad T_1 G$ gets mapped to the zero-vector field. We conclude that Θ is a vector space isomorphism (its inverse is obviously evaluation in $\mathbf{1}$). Note that $V\ (G)$ carries the subspace topology induced by $V(G)$. That $[\cdot, \cdot]$ is a Lie bracket on $T_1 G$ follows directly by trivial computations since $V\ (G)$ is a Lie algebra. □

We will in the rest of this section sometimes denote the left invariant vector field associated to an element $v \quad T_1 G$ by $v\ $ to avoid the clumsy notation $\Theta(v)$.

3.1.3 Remark (Left-right confusion) The reader may wonder now, why one uses left-invariant vector fields to compute the Lie algebra. Instead one could as well use *right-invariant vector fields*, i.e. $X(g) = T\rho_g X(\mathbf{1})$. This would also lead to a Lie algebra structure on $T_1 G$, however the induced Lie bracket would have the opposite sign (see Exercise 3.2 7.). Indeed there is no reason to prefer left-invariant vector fields over right-invariant ones, the choice for left-invariant fields is historically motivated and customary.

3.1.4 Example Consider the abelian Lie group $(\mathrm{R}^d, +)$. Since $T\mathrm{R}^d = \mathrm{R}^d \times \mathrm{R}^d$ we have $\mathbf{L}(\mathrm{R}^d) = T_0\mathrm{R}^n = \mathrm{R}^d$. Now $T\lambda_a(x, v) = (\lambda_a(x), \mathbf{d}\lambda_a(x)(v)) = (x + a, v)$, the equation $T\lambda_a \circ X = X \circ \lambda_a, a \quad \mathrm{R}^n$ yields that every left invariant vector field on R^n must be constant. The Lie bracket of constant vector fields is always 0. Thus the Lie algebra of the abelian Lie group R^d is an abelian Lie algebra R^n.

3.1.5 Example Consider the linear Lie group $\mathrm{GL}_n(\mathsf{K})$. Since $\mathrm{GL}_n(\mathsf{K})$ is an open subset of $M_n(\mathsf{K})$ we have $T_{I_n}\,\mathrm{GL}_n(\mathsf{K}) = M_n(\mathsf{K})$. The left invariant vector field $A\ $ corresponding to a matrix $A \quad M_n(\mathsf{K})$ is given by

$$A\ (X) = T_{I_n}\lambda_X(A) = (A, XA)$$

because $\lambda_X A = XA$ is a linear map whose derivative coincides with the left multiplication. Abusing notation and writing $A\ $ now for the principal part, we find with Equation (2.5)

$$\begin{aligned}[A, B] &= [A\ , B\ ](I_n) = \mathbf{d}B\ (I_n)A\ (I_n) - \mathbf{d}A\ (I_n)B\ (I_n) \\ &= \mathbf{d}B\ (I_n)A - \mathbf{d}A\ (I_n)B = AB - BA\end{aligned}$$

Therefore the Lie bracket on $\mathbf{L}(\mathrm{GL}_n(\mathsf{K})) = T_{I_n}\,\mathrm{GL}_n(\mathsf{K}) = M_n(\mathsf{K})$ is given by the commutator bracket and we recover exactly the Lie algebra we computed in Example 1.6.1.

The observation in Example 3.1.5 will be true for all linear Lie groups where the new definition of the Lie algebra coincides with the old. Before we can prove this, we will have to introduce the Lie group expoinential in the next section though. We now associate to every Lie group morphism a morphism of Lie algebras.

3.1.6 Lemma *If $f : G \to H$ is a Lie group morphism (i.e. a smooth group homomorphism) then the map $\mathbf{L}(f) := T_1 f : \mathbf{L}(G) \to \mathbf{L}(H)$ is a morphism of Lie algebras, i.e. $\mathbf{L}(f)([v, w]) = [\mathbf{L}(f)(v), \mathbf{L}(f)(w)],\ \ v, w \quad \mathbf{L}(G)$.*

Proof. Let $v \quad T_1G$ and $\tilde{v} := T_1 f(v) \quad T_1H$. Since f is a group morphism we have

$$\tilde{v}\ (f(g)) = T\lambda_{f(g)}T_1 f(v) = T(\lambda_{f(g)} \circ f)(v) = Tf(T_1\lambda_g(v)) = Tf(v\ (g)). \tag{3.4}$$

Hence for every $v \quad T_1G$ the left-invariant vector field $v\ $ is f-related to $\tilde{v}\ $. As f-relatedness is inherited by the Lie bracket of vector fields (Exercise 2.3 6.) we see that

$$Tf \circ [v\ , w\ ] = [\tilde{v}\ , \tilde{w}\ ] \circ f$$

evaluating in $\mathbf{1}_G$, we obtain the claimed formula since $f(\mathbf{1}_G) = \mathbf{1}_H$. □

3.1.7 Remark We defined a Lie algebra morphism $\mathbf{L}(\phi)$ for a continuous group morphism $\phi\colon G_1 \to G_2$ between **linear** Lie groups in Proposition 1.8.1. The map we defined coincides with the one we just introduced in Lemma 3.1.6, cf. Example 3.3.11.

3.1.8 Lemma *Let $\phi\colon G \to H$ be an isomorphism of Lie groups, then $\mathbf{L}(\phi)$ is an isomorphism of Lie algebras, and for each $x \in \mathbf{L}(G)$ the following holds*

$$\phi_* x^\ell := T\phi \circ x^\ell \circ \phi^{-1} = (\mathbf{L}(\phi)x)^\ell \tag{3.5}$$

Proof. If $\psi\colon H \to G$ is the inverse of ϕ, then $\phi \circ \psi = \mathrm{id}_H$ and $\psi \circ \phi = \mathrm{id}_G$. From Exercise 3.2 4. we deduce that $\mathbf{L}(\phi) \circ \mathbf{L}(\psi) = \mathrm{id}_{\mathbf{L}(H)}$ and $\mathbf{L}(\psi) \circ \mathbf{L}(\phi) = \mathrm{id}_{\mathbf{L}(G)}$. Now (3.5) is a direct consequence of (3.4). □

One may ask what the Lie bracket of a general Lie group captures of the structure of the underlying group. The answer is connected to the adjoint action.

3.1.9 Definition For a Lie group G, and $g \in G$, then the conjugation $c_g\colon G \to G, c_g(h) = ghg^{-1}$ is smooth. We define the *adjoint representation* of G as

$$\mathrm{Ad}\colon G \to \mathrm{GL}(\mathbf{L}(G)), \quad \mathrm{Ad}(g) = \mathrm{Ad}_g := \mathbf{L}(c_g).$$

With more work one can establish (cf. [HN12, Lemma 9.2.20]) the following:

3.1.10 Lemma *For a Lie group G we have* $\mathbf{L}(\mathrm{Ad}) = \mathrm{ad}\colon \mathbf{L}(G) \to \mathrm{Lin}(\mathbf{L}(G), \mathbf{L}(G))$, *where* $\mathrm{ad}(x)(y) = [x, y]$.

The proof of Lemma 3.1.10 is somewhat similar to the proof of Lemma 1.5.4 and requires the introduction of the so called Lie group exponential which we will treat now. In essence the Lie bracket of a general Lie group measures commutativity of the group.

3.2 Exercises

1. From the proof of Proposition 3.1.2. Show that $v_\ell : G \to TG, g \mapsto T_1\lambda_g(v)$ is a smooth left-invariant vector field for $v \in T_1G$.

2. Let $v \in \mathbf{L}(G)$ and denote by v_ℓ, v_r the left-(/right-)invariant vector field constructed from v. Show that v_ℓ and $-v_r$ are ι-related and deduce $[v_\ell, w_\ell] = -[v_r, w_r]$.

   Hint: By Exercise 3.1 5. b) $T_1\iota(v) = -v$.

3. Let G be an abelian Lie group. Show that the Lie bracket of $\mathbf{L}(G)$ is trivial, i.e. the Lie algebra is abelian.

   Hint: Since G is abelian all left-invariant vector fields are also right-invariant...

4. Let G_1, G_2, G_3 be Lie groups and $\phi\colon G_1 \to G_2$ and $\psi\colon G_2 \to G_2$ be morphisms of Lie groups. Prove that $\mathbf{L}(\mathrm{id}_{G_1}) = \mathrm{id}_{\mathbf{L}(G_1)}$ and $\mathbf{L}(\psi \circ \phi) = \mathbf{L}(\psi) \circ \mathbf{L}(\phi)$

3.2. The Lie group exponential

We generalise the matrix exponential to arbitrary Lie groups.

3.2.1 Lemma *Let G be a Lie group. Each left invariant vector field X on G is complete.*

Proof. Let $g \in G$ and $\gamma\colon I \to G$ be the unique maximal integral curve of $X \in V(M)$, cf. Theorem 2.3.6 with $\gamma(0) = g$. For each $h \in G$ we have $T\lambda_h(X) = X \circ \lambda_h$ which implies that $\eta := \lambda_h \circ \gamma$ is also an integral curve of X. Put $h = \gamma(s)g^{-1}$ for some $s > 0$. Then

$$\eta(0) = (\lambda_h \circ \gamma)(0) = h\gamma(0) = hg = \gamma(s),$$

and the uniqueness of integral curves implies that $\gamma(t+s) = \eta(t)$ for all t in the interval $I \cap (I - s)$ which is nonempty because it contains 0. In view of the maximaloty of I, it follows that $I - s \subseteq I$, and hence that $I - ns \subseteq I$ for all $n \in \mathbb{N}$, so the interval I is unbounded from below. Applying the same argument for an $s < 0$, we see that I is also unbounded from above. Hence $I = \mathbb{R}$ which means that I is complete. □

3.2.2 Definition Define the *Lie group exponential* for a Lie group G as

$$\exp_G\colon \mathbf{L}(G) \to G, \quad \exp_G(x) := \gamma_x(1),$$

where $\gamma_x\colon \mathbb{R} \to G$ is the unique maximal integral curve of the left invariant vector field x , satisfying $\gamma_x(0) = \mathbf{1}$. This means that γ_x is the unique solution of the initial value problem

$$\gamma(0) = \mathbf{1}, \quad \frac{d}{dt}\gamma(t) = x (\gamma(t)) = T_{\mathbf{1}}\lambda_{\gamma(t)}(x), \quad t \in \mathbb{R}.$$

3.2.3 Example For the linear Lie group $\mathrm{GL}_n(\mathbb{K})$ we saw in Proposition 1.4.6 that the smooth function $\phi_X(t) = e^{tX}$ is the integral curve of the right invariant vector field $X^r(A) = XA$ for $X \in M_n(\mathbb{K})$ and $A \in \mathrm{GL}_n(\mathbb{K})$. However, we also saw in Exercise 1.4 4. that it is also the integral curve of the left invariant vector field $X (A) = AX$. Since all left invariant vector fields on $\mathrm{GL}_n(\mathbb{K})$ are of this form by Example 3.1.5, we deduce that the Lie group exponential of $\mathrm{GL}_n(\mathbb{K})$ is simply the matrix exponential.

For the matrix exponential we had a lot of information. Our aim is now to establish the same statements for the Lie group exponential of an arbitrary Lie group.

3.2.4 Lemma *Let G be a Lie group. For each $x \in \mathbf{L}(G)$,*

(a) *the curve $\gamma_x\colon \mathbb{R} \to G$, $\gamma_x(t) = \exp_G(tx)$ is a smooth homomorphism of Lie groups with $\dot\gamma_x(0) = x$.*

(b) *The global flow of the left invariant vector field x is given by*

$$\mathrm{Fl}^{x}(t,g) = g\gamma_x(t) = g\exp_G(tx).$$

(c) *If $\gamma\colon \mathbb{R} \to G$ is a smooth homomorphism of Lie groups and $\dot\gamma(0) = x$, then $\gamma = \gamma_x$. In particular, the map $\mathrm{Hom}(\mathbb{R}, G) \to \mathbf{L}(G), \gamma \to \dot\gamma(0)$ is a bijection (here $\mathrm{Hom}(\mathbb{R}, G)$ denotes the set of smooth homomorphisms of Lie groups $\mathbb{R} \to G$.*

Proof. (a), (b) Since γ_x is an integral curve of the smooth vector field x, it is a smooth curve. By smoothness of multiplication in the Lie group G, the map $\mathrm{Fl}^{x}(t,g) = g\gamma_x(t)$ defines a smooth map $\mathbb{R} \times G \to G$. For a left invariant vector field x we have for each $g \in G$ and $\mathrm{Fl}^{x}_g(t) := \mathrm{Fl}^{x}(t,g)$ the relation

$$(\dot{\mathrm{Fl}}^{x}_g)(t) = T(\lambda_g)\dot\gamma_x(t) = T(\lambda_g)x(\gamma_x(t)) = x(g\gamma_x(t)) = x(\mathrm{Fl}^{x}_g(t)).$$

Therefore, Fl^{x}_g is an integral curve for x with $\mathrm{Fl}^{x}_g(0) = g$ and this proves that Fl^{x} is the unique maximal flow of the complete vector field x. In particular for $t, s \in \mathbb{R}$ we obtain

$$\gamma_x(t+s) = \mathrm{Fl}^{x}(t+s, \mathbf{1}_G) = \mathrm{Fl}^{x}(t, \mathrm{Fl}^{x}(s, \mathbf{1}_G)) = \mathrm{Fl}^{x}(s, \mathbf{1}_G)\gamma_x(t) = \gamma_x(s)\gamma_x(t).$$

Hence γ_x is a group homomorphism $(\mathbb{R}, +) \to G$.

(c) If $\gamma\colon (\mathbb{R}, +) \to G$ is a smooth group homomorphism, then $\mathrm{Fl}^{x}(t,g) := g\gamma(t)$ defines a global flow on G whose infinitesimal generator is the vector field

$$X(g) = \frac{d}{dt}\Big|_{t=0} Fl^{x}(t,g) = T(\lambda_g)\dot\gamma(0).$$

We conclude that $X = x$ for $x = \dot\gamma(0)$, so that X is a left invariant vector field. Since γ is the unique integral curve through $\mathbf{1}_G$, it follows that $\gamma = \gamma_x$. In view of (a), this proves (c). □

3.2.5 Proposition (Naturality of the Lie group exponential) *Let $\phi\colon G_1 \to G_2$ be a morphism of Lie groups and $\mathbf{L}(\phi)\colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ its differential at $\mathbf{1}$. Then*

$$\exp_{G_2} \circ \mathbf{L}(\phi) = \phi \circ \exp_{G_1}. \tag{3.6}$$

Proof. For $x \in \mathbf{L}(G_1)$ we consider the smooth homomorphism

$$\gamma_x \in \mathrm{Hom}(\mathbb{R}, G_1), \quad \gamma_x(t) = \exp_{G_1}(tx).$$

Following Lemma 3.2.4 we have

$$\phi \circ \gamma_x(t) = \exp_{G_2}(ty)$$

for $y = (\phi \circ \gamma_x)^{\cdot}(0) = \mathbf{L}(\phi)x$, because $\phi \circ \gamma_x \colon \mathbb{R} \to G_2$ is a smooth group homomorphism. For $t = 1$, we obtain in particular

$$\exp_{G_2}(\mathbf{L}(\phi)x) = \phi(\exp_{G_1}(x)),$$

which was all we had to show. □

3.2.6 Proposition *For a Lie group G, the Lie group exponential $\exp_G \colon \mathbf{L}(G) \to G$ is smooth and satisfies $T_0 \exp_G = \mathrm{id}_{\mathbf{L}(G)}$. In particular, $\exp_G$ is a local diffeomorphism at 0 in the sense that it maps some 0-neighborhood in $\mathbf{L}(G)$ diffeomorphically onto some $\mathbf{1}_G$-neighborhood in G.*

Proof. Let $n \in \mathbb{N}$. From Lemma 3.2.4 (c) we deduce that

$$\exp_G(nx) = \gamma_x(n) = \gamma_x(1)^n = \exp_G(x)^n \tag{3.7}$$

for $x \in \mathbf{L}(G)$. Since the n-fold multiplication map $G^n \to G, (g_1, \ldots, g_n) \mapsto g_1 \cdot g_2 \cdots g_n$ is smooth, the nth power map $G \to G, g \mapsto g^n$ is smooth. Therefore it suffices to verify smoothness of $\exp_G$ in some 0-neighborhood W. Then (3.7) implies smoothness on nW for each n and hence on all of $\mathbf{L}(G)$.

Now it suffices to notice that the map $\Psi \colon \mathbf{L}(G) \to \mathcal{V}(G), x \mapsto x$ satisfies the assumptions of Proposition A.3.13 since the map

$$\mathbf{L}(G) \times G \to TG, (x, g) \mapsto x(g) = gx$$

is smooth.[1] So

$$\Phi \colon \mathbb{R} \times \mathbf{L}(G) \times G \to G, (t, x, g) \mapsto g\gamma_x(t) = g \exp_G(tx)$$

is smooth on a neighborhood of $(0, 0, \mathbf{1}_G)$. In particular, for some $t > 0$, the map $x \mapsto \exp_G(tx)$ is smooth on a 0-neighborhood of $\mathbf{L}(G)$ and this proves that $\exp_G$ is smooth.

Finally, we observe that

$$T_0 \exp_G(x) = \left.\frac{d}{dt}\right|_{t=0} \exp_G(tx) = \dot\gamma_x(0) = x.$$

In other words $T_0 \exp_G = \mathrm{id}_{\mathbf{L}(G)}$. □

Since $\exp_G$ is locally invertible near 0, we can thus construct for every Lie group G its inverse on an open identity neighborhood which is usually dubbed the Lie group logarithm and abbreviated log. Note that log exists only on some identity neighborhood.

[1] Actually we cheat here a bit and suppress rewriting everything in charts as we should. The reader is invited to do it herself and check the details!

3.2.7 Lemma *If $x, y \in \mathbf{L}(G)$ commute, i.e. $[x, y] = 0$, then*

$$\exp_G(x + y) = \exp_G(x)\exp_G(y).$$

Proof. If x, y commute, then the associated left invariant vector fields x^{λ}, y^{λ} also satisfy $[x^{\lambda}, y^{\lambda}] = 0$. Left invariant vector fields are complete by Lemma 3.2.1. Hence their flows $\mathrm{Fl}^X, \mathrm{Fl}^Y$ commute by Lemma 2.3.15. In other words for all $t, s \in \mathbb{R}$ we have

$$\exp_G(tx)\exp_G(sy) = \exp_G(sy)\exp(tx).$$

Therefore $\gamma(t) := \exp_G(tx)\exp_G(ty)$ is a smooth group homomorphism with $\dot{\gamma}(0) = T_{1,1}(m_G)(x, y) = x+y$ (see Lemma 3.0.9). By Lemma 3.2.4 this leads to $\gamma(t) = \exp(t(x+y))$, and for $t = 1$ we obtain the statement of the Lemma. □

In the general case that elements do not commute we still have the Lie-Trotter formula (note that the proof of Proposition 1.4.13 carries over exactly as stated if we replace the matrix exponential with the Lie group exponential and use the properties already established in this section:

3.2.8 Proposition ((Lie-)Trotter product formula (for general Lie groups)) *Let G be a Lie group and $x, y \in \mathbf{L}(G)$ then*

$$\exp_G(x + y) = \lim_{k\to\infty}\left(\exp_G\left(\frac{x}{k}\right)\exp_G\left(\frac{y}{k}\right)\right)^k.$$

3.2.9 Lemma (Canonical coordinates) *Let G be a Lie group and $b_1, \ldots, b_n$ be a basis for $\mathbf{L}(G)$. Then the following maps restrict to diffeomorphisms of some 0-neighborhood in $\mathbb{R}^n$ to some open $\mathbf{1}_G$-neighborhood in G:*

(i) $x \mapsto \exp_G(x_1b_1 + \cdots + x_nb_n)$ *(coordinates of the first kind)*

(ii) $x \mapsto \exp_G(x_1b_1)\cdots\exp_G(x_nb_n)$ *(coordinates of the second kind).*

Proof. (i) is an immediate consequence of Proposition 3.2.6. For (ii) we use that $T_0\exp_G = \mathrm{id}_{\mathbf{L}(G)}$ and by Lemma 3.0.9 we have $T_{(\mathbf{1},\mathbf{1})}m_G(x, y) = x + y$. Therefore

$$\Xi\colon \mathbb{R}^n \to G, \quad \Xi(x) = \exp_G(x_1b_1)\cdots\exp_G(x_nb_n)$$

satisfies $T_0\Xi(X) = \sum_{i=1}^n x_ib_i$. The claim thus follows from the Inverse function Theorem A.2.7. □

Canonical coordinates are often a convenient way to describe the Lie group locally around the identity. They are one reason physicists often talk about "the Lie group" while working in reality with a Lie algebra (and then say that they could always get the group by "exponentiating the algebra"). We note that canonical coordinates have many uses, not the least in numerical analysis to construct numerical methods for Lie groups. The interested reader could consult for example [OM01].

3.3 Exercises

1. Show that for the abelian Lie group $(\mathbb{R}^d, +)$ the Lie group exponential is just the identity $\mathrm{id}_{\mathbb{R}^d}$.
   Hint: Check Example 3.1.4.

2. Let G, H be linear Lie groups and $\phi \colon G \to H$ a Lie group morphism. Show that
$$T_1\phi(x) = \frac{d}{dt}\Big|_{t=0} \phi \circ \exp(tx)$$
   Deduce that the two definitions of $\mathbf{L}(\phi)$ from Lemma 3.1.6 and Proposition 1.8.1 coincide.

3. We go back to the Heisenberg algebra $\mathfrak{h}$ from Exercise 1.9 1.
   a) Go back to the Exercise, recall the exponential of the Heisenberg algebra and prove that $\exp(\mathfrak{h})$ is isomorphic to the Heisenberg group from Example 1.1.13.
   b) Use the computations from Exercise 1.9 1. to give explicit formulae for canonical coordinates of first and second kind for the Heisenberg group.

3.3. Lie group morphisms and the closed subgroup theorem

In this section we study morphisms of Lie groups and Lie subgroups.

3.3.1 Lemma *Let G, H be Lie groups and $\phi \colon G \to H$ be a group homomorphism. Then the following are equivalent:*

(a) *ϕ is smooth in an identity neighborhood of G,*

(b) *ϕ is smooth.*

(c) *There exists a linear map $\psi \colon \mathbf{L}(G) \to \mathbf{L}(H)$ satisfying*
$$\exp_H \circ \psi = \phi \circ \exp_G \tag{3.8}$$

Proof. The equivalence of (a) and (b) was already proved in Exercise 3.1 2. Only (a) (b) is interesting: If ϕ is smooth on the identity neighborhood U we have with smoothness of the left translations for every $g \quad G$

$$\phi \circ \lambda_g(x) = \phi(gx) = \phi(g)\phi(x) = \lambda_{\phi(g)}\phi(x), \quad x \quad U.$$

Hence the smoothness of ϕ on U implies the smoothness of ϕ on gU, and therefore (b).

(b) (c) If ϕ is smooth, then $\psi = \mathbf{L}(\phi)$ satisfies (3.8).

(c) (a) If ψ is a linear map satisfying (3.8), then the fact that $\exp_G$ and $\exp_H$ are local diffeomorphisms by Proposition 3.2.6 and the smoothness of the linear map ψ imply (a). □

3.3.2 Proposition *For a morphism $\phi\colon G \to H$ of Lie groups the following holds:*

(a) $\ker \mathbf{L}(\phi) = \{x \quad \mathbf{L}(G)\colon \exp_G(\mathbb{R}x) \quad \ker \phi\}$

(b) *ϕ is an open map*[2] *if and only if $\mathbf{L}(\phi)$ is surjective.*

(c) *If $\mathbf{L}(\phi)$ is a linear isomorphism and ϕ is bijective, then ϕ is an isomorphism of Lie groups.*

Proof. (a) If $x \quad \ker \mathbf{L}(\phi)$ we have

$$\{\mathbf{1}_H\} = \exp_H(\mathbb{R}\mathbf{L}(\phi)x) = \phi(\exp_G(\mathbb{R}x)).$$

The converse inclusion uses Proposition 3.2.6 and is left as Exercise 3.4 2.

(b) Assume first that ϕ is an open map. Now $\exp_G, \exp_H$ are local diffeomorphisms such that

$$\exp_H \circ \mathbf{L}(\phi) = \phi \circ \exp_G \tag{3.9}$$

implies that there exists some open 0-neighborhood in $\mathbf{L}(G)$ on which $\mathbf{L}(\phi)$ is an open map, hence $\mathbf{L}(\phi)$ is surjective. If conversely, $\mathbf{L}(\phi)$ is surjective, then $\mathbf{L}(\phi)$ is an open map, so (3.9) implies that there exists an open $\mathbf{1}_G$-neighborhood U such that $\phi|_U$ is an open map. If $O \quad G$ is open we can pick for each $g \quad O$ an open $\mathbf{1}_G$-neighborhood $U_g \circ U$ such that $gU_g \quad O$. Then

$$\phi(O) \quad \phi(gU_g) = \phi(g)\phi(U_g) = \lambda_{\phi(g)}(\phi(U_g)).$$

As $U_g \quad U$, the set on the right hand side is open and contains $\phi(g)$. So $\phi(O)$ is a neighborhood for every $g \quad O$, whence it is open. We deduce that ϕ is an open map.

(c) Since ϕ is bijective, from (3.9) we derive that the group homomorphism ϕ^{-1} satisfies

$$\phi^{-1} \circ \exp_H = \exp_G \circ \mathbf{L}(\phi)^{-1}.$$

[2] A map is called open if it maps open sets to open sets.

Since $\exp_H$ is smoothly invertible on some open identity neighborhood, Lemma 3.3.1 shows that also ϕ^{-1} is smooth. □

We slightly extend Definition 1.4.1 so that it makes sense for general Lie groups.

3.3.3 Definition Let G be a topological group. We call a continuous group morphism $(\mathrm{R}, +) \to G$ a 1*-parameter (sub-)group*.

With the tools available the proof of Theorem 1.4.11 generalises to the following:

3.3.4 Theorem (1-parameter Group Theorem) *Let G be a Lie group. Then every one parameter group $\gamma\colon \mathrm{R} \to G$ is of the form $\gamma_x(t) = \exp_G(tx)$, where $x \quad \mathbf{L}(G)$ is the element $x = \dot\gamma(0)$. So in particular, every* 1*-parameter group defines a Lie group morphism.*

As a consequence we obtain.

3.3.5 Theorem (Automatic Smoothness Theorem) *Let G, H be Lie groups and $\phi\colon G \to H$ a continuous group morphism. Then ϕ is smooth.*

Proof. Arguing exactly as in the first part of the proof of Proposition 1.8.1 we obtain a linear map $\mathbf{L}(\phi)\colon \mathbf{L}(G) \to \mathbf{L}(H)$ with $\phi \circ \exp_G = \exp_H \circ \mathbf{L}(\phi)$. Now ϕ is smooth by Lemma 3.3.1. □

3.3.6 Corollary *A topological group G carries at most one Lie group structure.*

Proof. If $\phi\colon G_1 \to G_2$ is a topological group isomorphism of two Lie groups, the Automatic Smoothness Theorem 3.3.5 shows that ϕ and its inverse ϕ^{-1} are smooth, whence isomorphisms of Lie groups. □

The Closed Subgroup Theorem

With the knowledge about Lie group morphisms we can now obtain useful information about Lie subgroups.

3.3.7 Proposition *Let $H \quad G$ be a Lie subgroup of a Lie group G. Then H is a closed subgroup of G.*

Proof. Since H is a Lie subgroup, the inclusion $i\colon H \to G$ is smooth by Lemma 2.1.16. As i is an injective Lie group morphism, Proposition 3.3.2 implies that $\mathbf{L}(i)\colon \mathbf{L}(H) \to \mathbf{L}(G)$ is injective, whence $\mathbf{L}(i)$ identifies the Lie algebra of H with the subspace $F := \mathbf{L}(i)\mathbf{L}(H)$ of $\mathbf{L}(G)$. Pick $0 \in V \subseteq^{\circ} \mathbf{L}(G)$ such that $\exp_G|_V$ is a diffeomorphism onto an open subset of G. Taking $\varepsilon > 0$ small enough, the closed ball $A := \overline{B_\varepsilon(0)} \subseteq V$ is compact. Further $A \cap F$ is a closed subset of a compact set, whence compact. Now

- $K := \exp_H(\mathbf{L}(\imath)^{-1}(A \cap F))$ is a compact $\mathbf{1}_H$ neighborhood in H-[3]
- $W := \exp_G(B_\varepsilon(0))$ is an open $\mathbf{1}_G$-neighborhood in G.
- Identify H with $i(H)$ then $W \cap H \subseteq K$ due to naturality of the exponential maps.

The continuous map $\delta(u, v) = uv^{-1}$, cf. (B.1) takes $(\mathbf{1}_G, \mathbf{1}_G)$ to $\mathbf{1}_G \in W$, so there is $\mathbf{1}_G \in U \subseteq^{\circ} G$ with $\delta(U, U) \subseteq W$.

After these preparations, let us prove that if $(a_n)_{n \in \mathbb{N}} \subseteq H$ is a sequence with $a_n \to a$, then we must have $a \in H$ (i.e. we use that H is closed if every convergent sequence takes its limit in H). Define $b_n := a_n a^{-1} = \delta(a_n, a^{-1}), n \in \mathbb{N}$ and note that $(b_n)_n$ converges in G to $\mathbf{1}_G$. Picking N large enough we may assume that $b_n \in U$, $n \geq N$. Let us compute

$$W \ni \delta(b_n, b_N) = b_n(b_N)^{-1} = (a_n a^{-1})(a_N a^{-1})^{-1} = a_n a_N^{-1}, n \geq N.$$

Since $a_n \in H$ for every n, we see that for $n \geq N$ $a_n a_N^{-1} \in W \cap H \subseteq K$. Now K is compact. so we may replace $(a_n a_N^{-1})_n$ with a convergent subsequence $(a_{n_k} a_N^{-1})_k$ whose limit $c \in K \subseteq H$. However, this implies

$$a = \lim_{n \to \infty} a_n = \lim_{k \to \infty} a_{n_k} = (\lim_{k \to \infty} a_{n_k} a_N^{-1}) a_N = c a_N \in H.$$

This completes the proof. □

This yields an a posteriori justification as to why we only considered closed subgroups of $\mathrm{GL}_n(\mathbb{K})$ as linear Lie groups. Every Lie subgroup turns out to be closed anyway. However, we will prove something even more impressive: Every closed subgroup of a Lie group is indeed a Lie subgroup. We start with some preparation.

3.3.8 Definition Let G be a Lie group and $H \subseteq G$ a closed subgroup. Define the set

$$\mathbf{L}^e(H) := \{x \in \mathbf{L}(G)\colon \exp_G(\mathbb{R}x) \subseteq H\}$$

3.3.9 Lemma *For a closed subgroup H of a Lie group G, the set $\mathbf{L}^e(H)$ is a real vector subspace of $\mathbf{L}(G)$.*

[3] $\mathbf{L}(\imath)$ is a linear isomorphism onto F. Its inverse on F is thus continuous and the image of compact sets under continuous maps is compact.

Proof. By definition $\mathrm{R}\mathbf{L}^e(H) \subseteq \mathbf{L}^e(H)$, whence $\mathbf{L}^e(H)$ is closed under scalar multiplication. Since Proposition 3.2.8 establishes the Lie-Trotter product formula for G, the proof now works as in Lemma 1.5.2. □

3.3.10 Lemma *Let H be a Lie subgroup of G. Let $i\colon H \to G$ be the inclusion. Then we can identify $\mathbf{L}(H)$ with $\mathbf{L}(i)(\mathbf{L}(H)) \subseteq \mathbf{L}(G)$ and the following holds:*

(a) *The Lie group exponential $\exp_H$ of G restricts to the one of H.*

(b) *Moreover, $\mathbf{L}(i)(\mathbf{L}(H)) = \mathbf{L}^e(H)$, in particular, $\mathbf{L}^e(H)$ is a Lie subalgebra.*

Proof. (a) Apply Proposition 3.2.5 to $i\colon H \to G$, to get $\exp_G \circ \mathbf{L}(i) = i \circ \exp_H$. Now i is just an inclusion and by Proposition 3.3.2 we can identify $\mathbf{L}(H)$ with its image $\mathbf{L}(i)\mathbf{L}(H)$. By abuse of notation $\exp_G|_{\mathbf{L}(H)} = \exp_H$.

(b) We identify again $\mathbf{L}(H)$ with its image $\mathbf{L}(i)\mathbf{L}(H) \subseteq \mathbf{L}(G)$. If $x \in \mathbf{L}(H)$, then $\exp_G(tx) = \exp_H(tx) \in H$ by part (a). Thus $\mathbf{L}(H) \subseteq \mathbf{L}^e(H)$. Conversely if $x \in \mathbf{L}(G)$ suh that for all $t \in \mathrm{R}$ we have $\exp(tx) \in H$. Taking the derivative $\frac{d}{dt}\big|_{t=0} \exp_G(tx) = x$. But since the curve runs in the submanifold H we see that $x \in T_1 H = \mathbf{L}(H)$. This completes the proof since the image of the Lie algebra morphism $\mathbf{L}(i)$ is a Lie subalgebra. □

With more work (and methods we are not going to discuss), one can show that $\mathbf{L}^e(H)$ is even a Lie subalgebra for subgroups of a Lie group which are not closed. We are now ready to identify all the bits and pieces we defined for general Lie groups with our earlier definitions for linear Lie groups.

3.3.11 Example Let G be a linear Lie group, then by Proposition 3.0.7, G is a Lie subgroup of $\mathrm{GL}_n(\mathrm{K})$, whence in particular the inclusion $i_G\colon G \to \mathrm{GL}_n(\mathrm{K})$ is a morphism of Lie groups. Taking its derivative at I_n, $\mathbf{L}(i_G)\colon T_{I_n}G \to \mathbf{L}(\mathrm{GL}_n(\mathrm{K})) = M_n(\mathrm{K})$ is just the inclusion of a linear subspace into $M_n(\mathrm{K})$, so in other words $\mathbf{L}(i_G)X = X$. The Lie bracket on $\mathbf{L}(G)$ is thus just the Lie bracket of $\mathrm{GL}_n(\mathrm{K})$, i.e. the matrix commutator, Example 3.1.5 restricted to the subspace $\mathbf{L}(G)$.

We defined in Definition 1.5.1 the Lie algebra of the linear Lie group G as the set $\mathbf{L}^e(G)$ (but wrote then $\mathbf{L}(G)$ for this set). By Lemma 3.3.10 we see that $\mathbf{L}^e(G) = \mathbf{L}(G)$ so both definitions of the Lie algebra of a linear Lie group coincide. Further, the Lie group exponential of G is just the restriction of the matrix exponential to G. So Definition 1.7.1 of the Lie group exponential for a linear Lie group coincides with the definition of the Lie group exponential via flows of vector fields for linear Lie groups.

We now state the main technical result needed to endow closed subgroups with a Lie group structure. Fortunately, the proof is completely analogous to the proof of Proposition 1.7.3 with the knowledge we already have.

3.3.12 Lemma *Let H be a closed subgroup of a Lie group G, then there exists a 0-neighborhood $\Omega \circ \mathbf{L}(G)$ such that*

$$\exp_G(\Omega \cap \mathbf{L}^e(H)) = \exp_G(\Omega) \cap H.$$

3.3.13 Theorem (Closed Subgroup Theorem) *Let H be a closed subgroup of a Lie group G, then H is a Lie subgroup of G.*

Proof. By definition of a Lie subgroup, we only have to prove that H is a submanifold of G. Pick Ω as in Lemma 3.3.10. By Proposition 3.2.6 we may shrink Ω, such that without loss of generality $\exp_G|_\Omega \colon \Omega \to \exp_G(\Omega) \circ G$ is a diffeomorphism whose inverse we call log. As in the proof of Proposition 3.0.7 one can now show that these maps yield a submanifold chart around the identity and exploiting the group structure of H, also for H. □

3.3.14 Example Let us take a closer look at closed subgroups of the Lie group $(V, +)$, where V is a finite-dimensional vector space. From Exercise 3.3 1. we know that $\exp_V = \mathrm{id}_V$. Let H be a closed subgroup. Then

$$\mathbf{L}(H) = \{x \quad V : \mathrm{R}x \quad H\} \quad H$$

is the largest vector subspace contained in H. If $E \quad V$ is a vctor space complement for $\mathbf{L}(H)$, then $V = \mathbf{L}(H) \times E$ picking $U \circ \mathbf{L}(H)$ and $W \circ E$ small enough we can achieve that $U \times V \circ \Omega$ is a 0-neighborhood, for Ω as in Lemma 3.3.12. Note that by construction we must have then $W \cap H = \{0\}$, so $H \cap E$ is discrete[4] One can show [HN12, Exercise 9.3.4] that there exist linearly independent $e_1, \dots e_k \quad E$ with

$$E \cap H = \mathrm{Z}e_1 + \cdots + \mathrm{Z}e_k. \text{ In particular, we find}$$
$$H = \mathbf{L}(H) \times \mathrm{Z}^k = \mathrm{R}^d \times \mathrm{Z}^k \quad \text{for } d = \dim \mathbf{L}(H).$$

Note that the group is a direct product of the Lie algebra and the lattice Z^k with $k \leq \dim V - d$.

3.3.15 Example (Closed subgroups of the circle) Let $H \quad \mathrm{T} \quad (\mathrm{C}^\times, \cdot)$ be a proper closed subgroup (i.e. $H = \mathrm{T}$). Then $\mathbf{L}(H) = \mathbf{L}(\mathrm{T})$ implies $\dim H < \dim \mathrm{T} = 1$, so H is a discrete subgroup, whence finite as T is compact. Now the map

$$q \colon \mathrm{R} \to \mathrm{T}, \quad t \to e^{2\pi i t}$$

is a Lie group morphism, whence $q^{-1}(H)$ is a proper subgroup of R, whence cyclic by Example 3.3.14. Then $H = q(q^{-1}(H))$ is also cyclic and therefore, H is one of the groups

$$C_n = \{z \quad \mathrm{T} : z^n = 1\} \text{ of } n\text{th roots of unity.}$$

[4] A subset $D \quad \mathrm{R}^n$ is discrete if for every $x \quad D$ there exists an open ball $B_r(x)$ such that $B_r(x) \cap S = \{x\}$.

3.4 Exercises

1. In this section we have appealed several times to earlier work. Maybe its a good idea to check our work. So revisit the proofs and work out the details which need to change/generalise for the proofs of

a) Theorem 3.3.4, revisit for this the proof of Theorem 1.4.11,

b) Theorem 3.3.5, revisit the proof of Proposition 1.8.1 for this and check what needs to change).

c) Lemma 3.3.12 and check that it can be proved with the methods in Proposition 1.7.3.

d) for the proof of Theorem 3.3.13 check again the proof of Proposition 3.0.7.

2. Establish the details for part (a) of Proposition 3.3.2, i.e. show that for a Lie group morphism $\phi\colon G \to H$ we have $\mathbf{L}^e(\ker \phi) = \ker \mathbf{L}(\phi)$.
Hint: One inclusion was already established, for the other use Proposition 3.2.6.

3.4. The Lie theorems (without proof)

So far we have seen that every Lie group comes with an associated Lie algebra. Further if $\phi\colon G \to H$ is a Lie group morphism, then we can take the derivative to obtain the Lie algebra morphism $\mathbf{L}(\phi)\colon \mathbf{L}(G) \to \mathbf{L}(H)$. So Lie groups and their morphisms give rise to Lie algebras and their morphisms. Does every Lie algebra also come as the associated Lie algebra of a Lie group?

This is answered by the so called Lie theorems which we mention here due to their importance but will not prove in this lecture (as they require more techniques). It turns out that Lie algebra morphisms between Lie algebras integrate under certain conditions to Lie group morphisms.

3.4.1 Theorem (Lie II) *Let G be a connected and simply connected[5] Lie group, H a Lie group and $\psi\colon \mathbf{L}(G) \to \mathbf{L}(H)$ be a Lie algebra morphism. Then there exists a Lie group morphism $\phi\colon G \to H$ with $\mathbf{L}(\phi) = \psi$.*

The second Lie theorem is of fundamental importance in representation theory of Lie groups which studies representations, i.e. group homomorphism $\phi\colon G \to \mathrm{GL}_n(\mathbb{C})$. Due to the second Lie theorem every Lie group representation corresponds uniquely to a Lie

[5] A topological space X is called simply connected if every continuous loop is homotopic to the constant path. This means that for any $\gamma\colon [0,1] \to X$ is continuous with $\gamma(0) = x = \gamma(1)$, one needs to be able to find a continuous map (a homotopy) $h\colon [0,1]^2 \to X$, with $h(0,t) = x$, t $[0,1]$ and $h(1,t) = \gamma(t)$.

algebra representation $\mathbf{L}(\phi)\colon \mathbf{L}(G) \to \mathfrak{gl}_n(\mathbb{C})$ if the Lie group G is simply connected. Note that this is not a serious restriction as one can construct to every Lie group its simply connected cover, which turns out to again be a Lie group.

The third Lie theorem now answers the question which (finite-dimensional) Lie algebras are Lie algebras of Lie groups.

3.4.2 Theorem (Lie III) *For every finite dimensional Lie algebra $\mathfrak{g}$ there exists a connected Lie group G with $\mathbf{L}(G) = \mathfrak{g}$.*

3.4.3 Remark In case you are wondering what happened to Lie I: We omitted the first Lie theorem as it is a purely local statement which does not admit a global formulation on the Lie group.

So the upshot of this discussion is that (finite dimensional) Lie algebras arise from Lie groups and they recover a great deal of information about the Lie group. Note however, that some information is lost. We have already seen this in Example 1.8.9. Moreover, there are possibly many Lie groups which have the same Lie algebra. For example, in the chapter on linear groups, we saw that $\mathbf{L}(G) = \mathbf{L}(G_0)$ for the unit component G_0 of a linear Lie group G (the same argument works also for the unit component of a general Lie group, we leave it as an exercise).

Indeed with a bit more work, one can quantify the information lost when passing from a Lie group to its Lie algebra. This leads to covering theory of Lie groups (see e.g. [HN12]). We will not discuss this route in these notes.

3.5 Exercises

1. Define for a Lie group G the unit component

$$G_0 := \{g \quad G\colon \quad \gamma\colon [0,1] \to G, \text{ continuous with } \gamma(0) = \mathbf{1}_G, \gamma(1) = g\}$$

a) Prove that G_0 is normal subgroup,

b) State and prove an analogue of Proposition 1.7.6.

c) Deduce that G_0 is a Lie group and show that $\mathbf{L}(G_0) = \mathbf{L}(G)$.

4. Lie group actions

In this part we consider actions of Lie groups on manifolds. These will allow us to recast many of the concepts we encountered for 1-parameter groups and flows of vector fields in the language of Lie groups. From an application point of view, actions of Lie groups are usually why one is interested in Lie groups in the first place. There are many mathematical theories where Lie groups act on the objects of interest and form the symmetry groups for these objects. Examples of this are canonical transformations in Hamiltonian mechanics, isometries in Riemannian geometry or the already mentioned Lie group methods for ODEs. In all these examples the Lie groups act smoothly on smooth manifolds.

4.0.1 Definition Let M be a smooth manifold and G a Lie group. A *(smooth) action of G on M* is a smooth map

$$\sigma\colon G \times M \to M$$

with the following properties

(A1) $\sigma(\mathbf{1}, m) = m$ for all $m \in M$.

(A2) $\sigma(g_1, \sigma(g_2, m)) = \sigma(g_1 g_2, m)$ for $g_1, g_2 \in G$, $m \in M$.

We also write

$$g.m := \sigma(g, m), \quad \sigma_g(m) := \sigma(g, m) \quad \sigma^m(g) := \sigma(g, m).$$

The map σ^m is called the *orbit map*.

For each smooth action σ, the map

$$\hat{\sigma}\colon G \to \mathrm{Diff}(M), \quad g \mapsto \sigma_g$$

is a group homomorphism, and any homomorphism $\gamma\colon G \to \mathrm{Diff}(M)$ for which the map

$$\sigma_\gamma\colon G \times M \to M, \quad (g, m) \mapsto \gamma(g)(m)$$

is smooth defines a smooth action of G on M.

4.0.2 Remark What we call here an action is sometimes called a *left action*. Likewise one defines a *right action* as a smooth map $\sigma_R : M \times G \to M$ with

$$\sigma_R(m, \mathbf{1}) = m, \quad \sigma_R(\sigma_R(m, g_1), g_2) = \sigma_R(m, g_1 g_2).$$

For $m.g := \sigma(m, g)$, this takes the form

$$m.(g_1 g_2) = (m.g_1).g_2$$

of an associativity condition. If σ_R is a smooth right action of G on M, then

$$\sigma_L(g, m) := \sigma_R(m, g^{-1})$$

defines a smooth left action of G on M. Conversely if σ_L is a smooth left action, then $\sigma_R(m, g) := \sigma_L(g^{-1}, m)$ defines a smooth right action of G on M. This translation is one-to-one, so that we may pass freely from one type of action to the other.

4.0.3 Example If $X \in V(M)$ is a complete vector field and $\mathrm{Fl}^X : \mathbb{R} \times M \to M$ its global flow. The flow satisfies (cf. Definition A.3.9)

$$\mathrm{Fl}^X(t+s, m) = \mathrm{Fl}^X(t, \mathrm{Fl}^X(s, m)) \text{ and } \mathrm{Fl}^X(0, m) = m, \quad \forall s, t \in \mathbb{R}, m \in M.$$

Hence, Fl^X defines a smooth action of $G = (\mathbb{R}, +)$ on M.

4.0.4 Definition Let G be a Lie group, $\mathfrak{g}$ a Lie algebra and V a finite dimensional vector space. Then a ...

- Lie group morphism $\pi: G \to \mathrm{GL}(V)$ is called a *representation of G on V*.
- homomorphism of Lie algebras $\phi: \mathfrak{g} \to \mathfrak{gl}(V)$ is called *representation of $\mathfrak{g}$ on V*.

Any representation defines a Lie group action of G on V via

$$\sigma(g, v) := \pi(g)(v).$$

Thus representations are the same as *linear actions*, i.e. actions on vector spaces for which the σ_g are linear. An easy computation (see also Corollary 1.8.6) shows

4.0.5 Lemma *If $\phi: G \to \mathrm{GL}(V)$ is a representation of G, then $\mathbf{L}(\phi): \mathbf{L}(G) \to \mathfrak{gl}(V)$ is a representation of its Lie algebra. We call $\mathbf{L}(\phi)$ the* derived representation *of ϕ.*

4.0.6 Example The group $\mathrm{GL}_n(\mathbb{R})$ acts naturally on $\mathbb{R}^n$ via matrix vector multiplication:

$$\sigma: \mathrm{GL}_n(\mathbb{R}) \times \mathbb{R}^n \to \mathbb{R}^n, \quad (X, v) \mapsto Xv.$$

The same holds for every linear Lie group $G \quad \mathrm{GL}_n(\mathrm{R})$. Note that this action comes from the representation $\pi\colon G \to \mathrm{GL}_n(\mathrm{R}) = \mathrm{GL}(\mathrm{R}^n), \quad X \to X$.

We saw in Section 1.3 that the orthogonal group $\mathrm{O}_{n+1}(\mathrm{R})$ consists of isometries of R^{n+1} with respect to the euclidean inner product $\cdot, \cdot$. Since $\mathrm{S}^n = \{x \quad \mathrm{R}^{n+1}\colon \ x, x \ = 1\}$ is a submanifold, the canonical linear action of the orthogonal group restricts to a Lie group action

$$\mathrm{O}_{n+1}(\mathrm{R}) \times \mathrm{S}^n \to \mathrm{S}^n, (X, z) \to X(v).$$

4.0.7 Example Let G be a Lie group. Then multiplication $\sigma := m_G\colon G \times G \to G$ defines a Lie group action of G on itself. In this case, $\sigma_g = \lambda_g$ for $g \quad G$.

The multiplication defines a smooth right action of G on itself. The corresponding Lie group action is

$$\sigma\colon G \times G \to G, \quad (g, h) \to hg^{-1} \text{ with } \sigma_g = \rho_g^{-1}.$$

There is a third Lie group action of G on itself, the *conjugation action*

$$\sigma\colon G \times G \to G, \quad (g, h) \to c_g(h) := ghg^{-1} \text{ with } \sigma_g = c_g.$$

4.0.8 Proposition *Let $\sigma\colon G \times M \to M$ be a Lie group action. Then the map*

$$\dot\sigma\colon \mathbf{L}(G) \to V(M), \quad \dot\sigma(x)(m) := \mathbf{L}(\sigma)(x)(m) := -T_{\mathbf{1}}(\sigma^m)(x)$$

is a homomorphism of Lie algebras.

Proof. Observe that for each $x \quad \mathbf{L}(G)$ the map $\mathbf{L}(\sigma)(x)\colon M \to TM$ is smooth, with $\mathbf{L}(\sigma)(x)(m) \quad T_{\sigma(\mathbf{1},m)}(M) = T_m M$. Hence it is a smooth vector field on M.

Now we need to prove that $\dot\sigma$ is a homomorphism of Lie algebras. Pick $m \quad M$ and write

$$\phi^m := \sigma^m \circ \iota_G\colon G \to M, \quad \phi^m(g) = g^{-1}.m$$

for the reversed orbit map. Then

$$\phi^m(gh) = (gh)^{-1}.m = h^{-1}.(g^{-1}.m) = \phi^{g^{-1}.m}(h),$$

which we write as $\phi^m \circ \lambda_g = \phi^{g^{-1}.m}$. Differentiating at $\mathbf{1} \quad G$, we obtain for $x \quad \mathbf{L}(G)$:

$$\begin{aligned} T_g(\phi^m)x\ (g) &= T_g(\phi^m)T_{\mathbf{1}}(\lambda_g)x = T_{\mathbf{1}}(\phi^m \circ \lambda_g)x = T_{\mathbf{1}}(\phi^{g^{-1}.m})x \\ &= T_{\mathbf{1}}(\sigma^{g^{-1}.m})T_{\mathbf{1}}(\iota_G)x = -T_{\mathbf{1}}(\sigma^{\phi^m(g)})x = \dot\sigma(x)(\phi^m(g)). \end{aligned} \tag{4.1}$$

Thus the left invariant vector field x on G is ϕ^m-related to the vector field $\dot\sigma(x)$ on M. Now the Related Vector Field Lemma (Exercise 2.3 4.) implies that for $x, y \quad \mathbf{L}(G)$ the vector field $[x\ , y\ ]$ is ϕ^m-related to $[\dot\sigma(s), \dot\sigma(x)]$, which for each $m \quad M$ implies

$$\begin{aligned} \dot\sigma([x, y])(m) &= T_{\mathbf{1}}(\phi^m)[x, y]\ (\mathbf{1}) = T_{\mathbf{1}}(\phi^m)[x\ , y\ ](\mathbf{1}) \\ &= [\dot\sigma(x), \dot\sigma y](\phi^m(\mathbf{1})) = [\dot\sigma(x), \dot\sigma(y)](m) \end{aligned}$$

□

4.0.9 Example Let $\sigma\colon \mathbb{R} \times M \to M$ be a Lie group action. Then Proposition 4.0.8 yields a Lie algebra morphism $\dot{\sigma}\colon \mathbf{L}(\mathbb{R}) \to V(M)$. Now recall that $\mathbf{L}(\mathbb{R}) = \mathbb{R}$ and left invariant vector fields on $\mathbb{R}$ are constant. Hence (4.1) becomes for all $t \in \mathbb{R}$ and $m \in M$

$$\frac{d}{dt}\sigma^m(-t) = \frac{d}{dt}\phi^m(t) = T_t\phi^m(1)\ (t) = \dot{\sigma}(1)(\phi^m(t)) = \dot{\sigma}(1)(\sigma^m(-t))$$

We deduce that $\frac{d}{dt}\sigma^m(t) = \dot{\sigma}(1)(\sigma^m(t))$, whence σ^m is an integral curve and $\sigma = \mathrm{Fl}^{\dot{\sigma}(1)}$. One also calls the vector field $\dot{\sigma}(1)$ the *fundamental vector field* of the action. Thus all Lie group actions of $(\mathbb{R}, +)$ are flows of complete vector fields, cf. Example 4.0.3.

Observe that in the last example it was no accident that the vector field constructed to a Lie group action was complete. Indeed this holds generally by the observation that the integral curves of these vector fields coincide with the orbits of the group action. We collect this together with a formula describing the flow of a vector field.

4.0.10 Lemma *If $\sigma\colon G \times M \to M$ is a smooth action and $x \in \mathbf{L}(G)$, then the global flow of the vector field $\dot{\sigma}(x)$ is given by $\mathrm{Fl}^{\dot{\sigma}(x)}(t, m) = \exp_G(-tx).m$. In particular,*

$$\dot{\sigma}(x)(m) = \left.\frac{d}{dt}\right|_{t=0} \exp_G(-tx).m.$$

Proof. In the proof of Proposition 4.0.8 we established that

$$T_g(\phi^m)x\ (g) = \dot{\sigma}(x)(\phi^m(g))$$

for $\phi^m(g) = g^{-1}.m$. Applying Proposition 3.2.6 we obtain

$$\left.\frac{d}{dt}\right|_{t=0} \exp_G(-tx).m = T_{\mathbf{1}}(\phi^m)T_0(\exp_G)x = T_{\mathbf{1}}(\phi^m)x = \dot{\sigma}(x)(m).$$

This shows in particular that these curves give the flow and proves the statement. □

4.1 Exercises

1. Show that the adjoint representation of a Lie group G from Definition 3.1.9 is a representation of G.
2. Check the details of Example 4.0.6.
3. Compute $\dot{\sigma}$ for the Lie group actions in Example 4.0.6 and Example 4.0.7.

4.1. Orbits and stabilisers

In this section we study the structure of smooth actions of Lie groups with some obersvations on orbits and stabilisers.

4.1.1 Definition Let $\sigma\colon G \times M \to M$ be a Lie group action. For $m \in M$, the set

$$O_m := G.m := \{g.m : g \in G\} = \{\sigma(g,m) : g \in G\}$$

is called the *orbit of* m. The action is said to be *transitive* if there exists only one orbit, i.e., for $x, y \in M$, there exists $g \in G$ with $y = g.x$. We write $M/G := \{O_m : m \in M\}$ for the set of G-orbits on N

4.1.2 Remark If $\sigma\colon G \times M \to M$ is an action on M, then the orbits form a partition of M (Exercise). A subset $R \subseteq M$ is called a set of *representatives for the action* if each G-orbit in M meets R exactly once: $\forall x \in M, |R \cap O_x| = 1$.

4.1.3 Example Consider the action of the circle group

$$\mathbb{T} = \{z \in \mathbb{C}^\times : |z| = 1\}$$

on $\mathbb{C}$ by $\sigma\colon \mathbb{T} \times \mathbb{C} \to \mathbb{C}, z.w := zw$. The orbits of this action are concentric circles $O_w = \{zw : z \in \mathbb{T}\} = \{c \in \mathbb{C} : |c| = |w|\}$. A set of representatives is given by $R := \{r + 0\mathrm{i} : r \in [0, \infty[\}$.

4.1.4 Example For the action of $\mathrm{GL}_n(\mathbb{K})$ on $\mathbb{K}^n$ by matrix-vector multiplication (cf. Example 4.0.6) there are only two orbits

$$O_0 = \{0\} \quad \text{and} \quad O_x = \mathbb{K}^n \setminus \{0\}, \text{ for } x \neq 0.$$

The first orbit is clear. For the second we observe that every $x \in \mathbb{K}^n \setminus \{0\}$ can be complemented to a basis of $\mathbb{K}^n$, hence arises as the first column of an invertible matrix X. Then $Xe_1 = x$ implies that $O_x = O_{e_1}$. The statement on the orbits follows.

4.1.5 Example For the conjugation action of $\mathrm{GL}_n(\mathbb{K})$ on itself, the orbits are similarity classes of matrices $O_A = \{XAX^{-1} : X \in \mathrm{GL}_n(\mathbb{K})\}$.

4.1.6 Definition Let $\sigma\colon G \times M \to M$ be an action of the group G on M. For $m \in M$, the subset

$$G_m := \{g \in G : g.m = m\}$$

is called the *stabiliser of* m. For $g \in G$, we write

$$\mathrm{Fix}(g) := M^g := \{m \in M : g.m = m\}$$

for the *set of fixed points of* g in M. We then have

$$m \in M^g \Leftrightarrow g \in G_m.$$

For a subset $S \subseteq M$, we write

$$G_S := \bigcap_{m \in S} G_m = \{g \in G : (\forall m \in S)\, g.m = m\},$$

and for $H \subseteq G$ we write

$$M^H := \{m \in M : (\forall h \in H)\; h.m = m\}$$

for the set of points in M fixed by all points in H.

4.1.7 Lemma *For each Lie group action $\sigma \colon G \times M \to M$, the following holds:*

(a) *For each $m \in M$, the stabiliser G_m of m is a Lie subgroup of G.*

(b) *For $m \in M$ and $g \in G$, $G_{g.m} = gG_m g^{-1}$.*

(c) *If $S \subseteq M$ is a G-invariant subset*[1]*, then G_S is a normal subgroup of G.*

The normal subgroup G_M of all elements which act as the identity on M is called the effectivity kernel *of the action. It is the kernel of $G \to \mathrm{Diff}(M), g \mapsto \sigma_g$.*

Proof. (a) That G_m is a subgroup follows from the definition of a group action. This subgroup is closed as by continuity of $\sigma^m \colon G \to M$ since $G_m = (\sigma^m)^{-1}(\{m\})$. Hence the Closed Subgroup Theorem 3.3.13 implies that G_m is a Lie subgroup.

(b) If $h \in G_m$, then

$$ghg^{-1}.(g.m) = ghg^{-1}g.m = g.m,$$

hence $ghg^{-1} \in G_{g.m}$ and thus $gG_m g^{-1} \subseteq G_{g.m}$. Similarly, we get $g^{-1}G_{g.m}g \subseteq G_{g^{-1}g.m} = G_m$. Now (c) is a direct consequence of (b). □

4.1.8 Proposition *Let $\sigma \colon G \times M \to M$ be a Lie group action with $\dot\sigma \colon \mathbf{L}(G) \to \mathcal{V}(M)$ as in Proposition 4.0.8. Then*

(a) *$m \in M^G \Rightarrow \dot\sigma(x)(m) = 0, \forall x \in \mathbf{L}(G)$. The converse holds for connected G.*

(b) *If $\dot\sigma(\mathbf{L}(G))(m) = T_m M$, then the orbit O_m of m is open.*

[1]This means that $G.S \subseteq S$, where the set on the left hand side is given by all elements of the form $g.s, g \in G, s \in S$.

Proof. (a) Suppose first, that m is a fixed point and let $x \in \mathbf{L}(G)$. Then by Lemma 4.0.10

$$\dot{\sigma}(x)(m) = \frac{d}{dt}\Big|_{t=0} \exp_G(-tx).m = \frac{d}{dt}\Big|_{t=0} m = 0.$$

If conversely, all vector fields $\dot{\sigma}(x)$ vanish at m, then m is a fixed point for all flows generated by these vector fields, which leads to $\exp_G(x).m = m$ for each $x \in \mathbf{L}(G)$. In particular, the group generated by $\exp_G(\mathbf{L}(G))$ is the identity component G_0 of G (this follows from Exercise 3.5 b)), is contained in G_m. As G is connected, we get $G = G_m$.

(b) Since the linear map $-T_1\sigma^m : \mathbf{L}(G) \to T_m M, x \mapsto \dot{\sigma}(x)(m)$ is surjective, the Implicit Function Theorem A.2.8 implies that $G.m = \sigma^m(G)$ is a neighborhood of m (Exercise: Work out the details!). Since all maps σ_g are diffeomorphisms of M, $\sigma_g(G.m) = gG.m = G.m$ also is a neighborhood of $g.m$. We conclude that $O_m = G.m$ is open. □

This yields an identification of the Lie algebra of the stabiliser group.

4.1.9 Corollary *If $\sigma: G \times M \to M$ is a smooth action, then for each $m \in M$*

$$\mathbf{L}(G_m) = \{x \in \mathbf{L}(G) : \dot{\sigma}(x)(m) = 0\}.$$

Proposition 4.1.8 shows in particular that the orbit O_m is a submanifold if m is a fixed point (then the submanifold is 0-dimensional) and if O_m is open. Our next goal will be to show that orbits of smooth actions always carry a natural manifold structure.

4.2 Exercises

1. Let $\sigma: G \times M \to M$ be a group action. Show that the orbits form a partition of M.

2. Show that the following maps define group actions and determine their orbits by naming a representative for each orbit:

a) $\mathrm{O}_n(\mathbb{R}) \times \mathbb{R}^n \to \mathbb{R}^n, (X, v) \mapsto Xv$;

b) $\mathrm{O}_n(\mathbb{R}) \times (\mathbb{R}^n \times \mathbb{R}^n) \to \mathbb{R}^n \times \mathbb{R}^n, (X, (a, b)) \mapsto (Xa, Xb)$.

3. For a complex number $\lambda \in \mathbb{C}$ consider the smooth action

$$\sigma: \mathbb{R} \times \mathbb{C} \to \mathbb{C}, \quad \sigma(t, z) := e^{t\lambda} z.$$

a) Sketch the orbits of this action in dependence of λ.

b) Under which conditions are there compact orbits?

c) Describe the corresponding fundamental vector field $\dot{\sigma}(1)$.

4.2. Homogeneous spaces

The main result[2] will be that for any Lie group action, all orbits carry a natural manifold structure. First, we take a closer look at transitive actions, i.e. actions with a single orbit.

4.2.1 Definition Let G be a group and H a subgroup. We write

$$G/H := \{gH : g \in G\}$$

for the *set of left cosets of H* in G and $q_{G/H}: G \to G/H, g \to gh$ for the quotient map. Then

$$\sigma: G \times G/H \to G/H, \quad (g, xH) \to gxH$$

defines a transitive action of G on the set G/H.

4.2.2 Definition Let G be a group and $\sigma_i: G \times M_i \to M_i, i = 1, 2$ two actions of the group G on sets. A map $f: M_1 \to M_2$ is called *G-equivariant* if

$$f(g.m) = g.f(m) \quad \text{holds for all } g \in G, m \in M_1.$$

4.2.3 Remark Let $\sigma: G \times M \to M$ be an action of the group G on the set M. Fix $m \in M$. Then the orbit map $\sigma^m: G \to O_m \subseteq M, \sigma^m(g) = g.m$ factors through a bijective map

$$\overline{\sigma}^m: G/G_m \to O_m, \quad gG_m \to g.m$$

which is equivariant with respect to the G-actions on G/G_m and M.

What we learn from the remark is that if we want a manifold structure on orbits of smooth actions, it is a natural idea to define a manifold structure on the coset space G/H for closed subgroups of the Lie group G. So far, G/H is only a set, let us first construct a suitable topology on it.

4.2.4 Lemma (Quotient topology on the coset space) *Let H be a closed subgroup of a Lie group G, then we endow the coset space G/H with the quotient topology with respect to the map $q: G \to G/H, q(g) = gH$, i.e.*

$$U \subseteq G/H \text{ is open if and only if } q^{-1}(U) \circ G.$$

The quotient topology is a Hausdorff topology and q is a continuous and open map.

[2] If you prefer video lectures, a lot of the material and proofs from this chapter is covered in the video lectures here: https://www.youtube.com/playlist?list=PLwDBpozLuH8P-ymXik39IOvJalnuRc5OH.

Proof. Note first that the quotient topology is a topology as it is the final topology with respect to the family of maps $\{q\}$ (cf. Definition A.1.5).

Since for each $O \subseteq^\circ G$ the product OH is open in G (Exercise B.1 2.), the openness of $OH = q^{-1}(q(O))$ shows that $q(O)$ is open in G/H. Hence q sends open sets to open sets and is thus an open map.

We have to prove that the quotient topology is Hausdorff. For this, let $g_1, g_2 \in G$ with $g_1 H \neq g_2 H$, i.e. $g_1 \notin g_2 H$. As $g_2 H$ is a closed subset, Lemma B.0.3 yields an open neighborhood U of g_1, such that $U \cap g_2 H = \emptyset$. Then we pick a symmetric $\mathbf{1}_G$-neighborhood $U_1 \subseteq^\circ G$ such that $U_1^{-1} U_1 g_1 \subseteq U$. We note that by construction $U_1 g_1 H$ and $U_1 g_2 H$ are now disjoint open subsets of G, such that

$$q(U_1 g_1 H) = q(U_1 g) \text{ and } q(U_1 g_2 H) = q(g_2 H)$$

are disjoint open subsets of G/H, separating $g_1 H$ and $g_2 H$. Thus G/H is Hausdorff. □

The next Lemma is a supplement to the Closed Subgroup Theorem 3.3.13:

4.2.5 Lemma *Let H be a closed subgroup of a Lie group G. Assume that E is a vector space complement of $\mathbf{L}(H)$ in $\mathbf{L}(G)$, then there exists $0 \in V_E \subseteq^\circ E$ such that*

$$\phi\colon V_E \times H \to \exp_G(V_E)H, \quad (x, h) \mapsto \exp_G(x)h$$

is a diffeomorphism onto an open subset of G.

Proof. We may pick $0 \in U_E \subseteq^\circ E$ and $0 \in U_H \subseteq^\circ \mathbf{L}(H)$ such that $\exp_H(U_H) = \exp_G(U_E \times U_H) \cap H$ and $\exp_G$ is invertible on $U_E \times U_H$ (cf. Lemma 3.3.10). In particular, $\exp_G(U_E) \cap H = \{\mathbf{1}\}$. So by an argument similar to the construction of coordinates of the second kind, we may shrink U_E, U_H such that

$$\Psi\colon U_E \times \exp_G(U_H) \to G, \Psi(x, h) = \exp_G(x)h$$

is a diffeomorphism onto an open subset of G. By continuity of multipliation there exists $V_E = -V_E \subseteq^\circ U_E$ with $\exp_G(V_E)\exp_G(V_E) \subseteq \exp_G(U_E)\exp_G(U_H)$. We claim that $\phi\colon V_E \times H \to \exp(V_E)H, \phi(x, h) = \exp_G(x)h$ is a diffeomorphism onto an open subset. To see this note that $\phi(x, h_1 h_2) = \phi(x, h_1)h_2$. Since $T_{(x,\mathbf{1})}\phi = T_{(x,0)}\exp_G$ for $x \in V_E$ is invertible, $T_{(x,h)}\phi = T_{\phi(x,\mathbf{1})}\rho_h T_{(x,\mathbf{1})}\phi$ is invertible for each $(x, h) \in V_E \times H$. So ϕ is a local diffeomorphism at each point. If we can show that ϕ is injective it will be a diffeomorphism (as claimed). Assume that $\phi(x, h) = \phi(x', h')$. Then

$$\begin{aligned}\exp_G(x)^{-1}\exp_G(x') &= h(h')^{-1} \\ &\in \exp_G(V_E)^2 \cap H \subseteq (\exp_G(U_E)\exp_G(U_H)) \cap H = \exp_G(U_H)\end{aligned}$$

Thus $\exp_G(x') \in \exp_G(x)\exp_G(U_H)$, so by injectivity of Ψ we must have $x = x'$ and this implies $h = h'$. □

4.2.6 Theorem *Let G be a Lie group, H a closed subgroup. Then the coset space G/H, endowed with the quotient topology, carries a natural manifold structure for which the quotient map $q\colon G \to G/H, q(g) = gH$ is a submersion.*

Moreover, $\sigma\colon G \times G/H \to G/H, (g, xH) \to gxH$ defines a smooth action of G on G/H.

Proof. As preparation choose a vector space complement E of $\mathbf{L}(H)$ in $\mathbf{L}(G)$. Then Lemma 3.3.12 implies that there is $(0,0) \quad V_E \times \Omega \circ E \times \mathbf{L}(H) = \mathbf{L}(G)$ such that $\exp|_{V_E \times \Omega}$ is a diffeomorphism onto an open $\mathbf{1}_G$-neighborhood and such that $\exp_G(V_E) \cap H = \{\mathbf{1}_H\}$. Shrinking V_E we deduce from Lemma 4.2.5 that

$$\phi\colon V_E \times H \to \exp_G(V_E)H, (x,h) \to \exp_G(x)h$$

is a homemorphism. Now observe that the following diagram commutes

$$\begin{array}{ccc} G \times G & \xrightarrow{m_G} & G \\ \downarrow{\scriptstyle \mathrm{id}_G \times q} & & \downarrow{\scriptstyle q} \\ G \times G/H & \xrightarrow{\sigma} & G/H \end{array} \tag{4.2}$$

This diagram will allow us to establish smoothness of σ, but first we need an intermediate step:

Step 1: σ is continuous As G/H is endowed with the quotient topology (see Lemma 4.2.4), the map q is a topological quotient map (more on this in Exercise 4.3 1.). Moreover G is locally compact as a finite dimensional manifold, whence $\mathrm{id}_G \times q\colon G \times G \to G \times G/H$ is a quotient map by Whiteheads Lemma (Exercise 4.3 2.). The vertical arrows in (4.2) are now quotient maps, whence σ inherits continuity from m_G by Exercise 4.3 1. b).

Step 2: The atlas of G/H: Let $W := q(\exp_G(V_E))$ and define a smooth map

$$p_E : q^{-1}(W) = \exp_G(V_E)H \to V_E, \quad \exp_G(x)h \to x.$$

Since $q^{-1}(W) = \exp_G(V_E)H$ is open in G, W is open in G/H. Moreover, $O \quad W$ is open if and only if $q^{-1}(O) \circ q^{-1}(W)$. Since $q^{-1}(O) = \exp_G(p_E(q^{-1}(O)))H$, this is equivalent to $p_E(q^{-1}(O))$ being open in V_E. Therefore $\psi\colon W \to V_E, q(g) \to p_E(g)$ is a homeomorphims and (ψ, W) is a chart of G/H.

For $g \quad G$, set $W_g := g.W$ and $\psi_g(x) := \psi(g^{-1}.x)$. Since all maps $\sigma_g\colon G/H \to G/H$ are homeomorphsims, we obtain an atlas $(\psi_g, W_g)_{g \quad G}$. To see that this atlas is smooth, let $g_1, g_2 \quad G$ and assume that $W_{g_1} \cap W_{g_2} = \quad$. We then have for $x \quad V_E$

$$\begin{aligned} \psi_{g_1} \circ \psi_{g_2}^{-1}(x) &= \psi(g_1^{-1}g_2.\psi^{-1}(x)) = \psi(g^{-1}g_2.q(\exp_G(x))) \\ &= \psi(q(g_1^{-1}g_2 \exp_G(x))) = p_E(g_1^{-1}g_2 \exp_G(x)) \end{aligned}$$

Since p_E is smooth, this map is smooth on its open domain, which shows that the change of chart maps are smooth and thus the atlas is smooth.

Step 3: Smoothness of the maps σ_g: For $g_1, g_2 \quad G$ we have $\sigma_{g_1}(W_{g_2}) = W_{g_1 g_2}$ and $\psi_{g_1 g_2} \circ \sigma_{g_1} = \psi_{g_2}$, which immediately implies that $\sigma_{g_1}|_{W_{g_2}} : W_{g_2} \to W_{g_1 g_2}$ is smooth. Since g_2 was arbitrary, all maps $\sigma_g, g \quad G$ are smooth. From $\sigma_g \circ \sigma_{g^{-1}} = \mathrm{id}_{G/H}$ we derive that they are diffeomorphisms.

Step 4: q is a submersion: The smoothness of q on $q^{-1}(W)$ follows from $\psi(q(g)) = p_E(g)$ and the smoothness of p_E on $q^{-1}(W)$. Moreover, $T_{\mathbf{1}_G H}\psi T_{\mathbf{1}_G} q = T_{\mathbf{1}_G} p_E : \mathbf{L}(G) \to E$ is the linear projection onto E with kernel $\mathbf{L}(H)$, hence surjective. This proves that $T_{\mathbf{1}} q$ is surjective. For each $g \quad G$ we have $q \circ \lambda_g = \sigma_g \circ q$, so Step 2 implies that q is smooth on all of G. Taking derivatives we obtain

$$T_g(q) \circ T_{\mathbf{1}_G}\lambda_g = T_{\mathbf{1}_G H}(\sigma_g) \circ T_{\mathbf{1}_G} q$$

and since all σ_g, λ_g are diffeomorphisms, this implies that all differentials $T_g q$ are surjective, whence q is a submersion.

Step 5: Smoothness of the action on G/H: Since q is a submersion, the product map $\mathrm{id}_G \times q : G \times G \to G \times G/H$ is also a submersion. Now the vertical arrows in (4.2) are submersions, whence commutativity of the diagram together with Proposition 2.2.11 implies smoothness of σ. □

4.2.7 Definition The manifolds of the form $M = G/H$ where H is a closed subgroup of a Lie group G, are called *homogeneous spaces*.

A consequence of the proof of Theorem 4.2.6 is worth recording as a corollary:

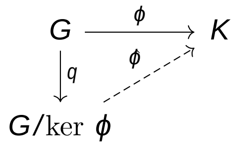
4.2.8 Corollary *Let G be a Lie group and H a closed subgroup. Then for any $x \quad G/H$ there exists $x \quad U \circ G/H$ and a smooth section $\sigma : U \to G$ for q such that*

$$m_\sigma : U \times H \to \sigma(U)H, \quad (u, h) \to \sigma(u)h$$

is a di eomorphism onto an open subset of G.

Consider the general setting of a Lie group morphism $\phi : G \to K$ and its kernel. We get a commutative diagram of (abstract) groups (shown to the right).

Since the kernel $\ker \phi$ is a closed normal subgroup, whence a normal Lie subgroup by the Closed Subgroup Theorem 3.3.13. Hence, q is a group morphism and a submersion with respect to the manifold structure from Theorem 4.2.6. As also $q \times q$ is a submersion we see that $m_{G/\ker \phi} \circ (q \times q) = q \circ m_G$ and $q \circ \iota_G = \iota_{G/\ker \phi} \circ q$ show that the group operations are smooth and $G/\ker \phi$ is a Lie group. (see also Exercise 4.3 4.)



$$\phi(q(x)) := \phi(x)$$

4.2.9 Example (The Heisenberg quotient) Let H be the Heisenberg group and Q be the Heisenberg quotient. In Exercise 3.1 6. we have seen that the map

$$\phi \colon H \to Q, \quad \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \to (x, y, e^{\mathrm{i}z})$$

is a surjective Lie group morphism whose kernel is a cyclic, discrete and normal subgroup D of H. In particular, $\bar{\phi} \colon H/D \to Q, xD \to \phi(x)$ is a group isomorphism.

Further, $\bar{\phi} \circ q = \phi$ is smooth, Proposition 2.2.11 shows that $\bar{\phi}$ is smooth, i.e. a Lie group morphism. Since $\bar{\phi}$ is injective, so is $\mathbf{L}(\bar{\phi})$. A direct computation shows that $T_{\mathbf{1}}\phi$ is surjective, so $T_{\mathbf{1}}\phi = T_{\mathbf{1}}\bar{\phi} \circ T_{\mathbf{1}}q$ shows that also $T_{\mathbf{1}}\bar{\phi}$ is surjective, whence $\mathbf{L}(\bar{\phi})$ is a linear isomorphism. By Proposition 3.3.2 the map $\bar{\phi}$ is an isomorphism of Lie groups.

The following corollary shows that for each smooth group action, all orbits carry natural manifold structures.

4.2.10 Corollary *Let $\sigma \colon G \times M \to M$ be a Lie group action. For each $m \in M$ the orbit map $\sigma^m \colon G \to M, g \to g.m$ factors through a smooth injective equivariant map*

$$\bar{\sigma}^m \colon G/G_m \to M, \quad gG_m \to g.m,$$

whose image is the set O_m.

Proof. The existence of $\bar{\sigma}^m$ is clear from Remark 4.2.3. Since the quotient map $q \colon G \to G/G_m$ is a submersion, the smoothness of $\bar{\sigma}^m$ follows from the smoothness of $\bar{\sigma}^m \circ q = \sigma^m$, Proposition 2.2.11. □

We can use the smooth map in Corollary 4.2.10 to endow every orbit with a manifold structure by declaring it a diffeomorphism. However, in general these manifold structures do not turn the orbits into submanifolds of the given manifold M on which G acts. We illustrate this in the next example.

4.2.11 Example (Torus action and dense wind) Let $\alpha \in \mathbb{R}$ and $\mathbb{T}^2$ the 2-dimensional torus from Example 3.0.4. Then we consider the action

$$\sigma_\alpha \colon \mathbb{R} \times \mathbb{T}^2 \to \mathbb{T}^2, \quad \sigma_\alpha(t, (e^{2\pi \mathrm{i} r}, e^{2\pi \mathrm{i} s})) := (e^{2\pi(r+\alpha t)\mathrm{i}}, e^{2\pi(s+t)\mathrm{i}}).$$

Consider now the orbit through $m = (1, 1) \in \mathbb{T}^2$.

If $\alpha \in \mathbb{Q}$ is rational, the orbit is a closed and periodic curve. In particular, we obtain a 1-dimensional closed Lie subgroup of the torus. By the Closed Subgroup Theorem 3.3.13 this orbit is a submanifold of $\mathbb{T}^2$.

However, if $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ is irrational, the orbit is not closed since the one parameter group $\sigma_\alpha^m : \mathbb{R} \to \mathbb{T}^2$ is injective (and while $\mathbb{T}^2$ is compact, $\mathbb{R}$ is not!). Hence the closure of σ_α^m is a closed subgroup of dimension at least 2, which shows that the orbit is dense in $\mathbb{T}^2$. We call this orbit the ***dense wind***. It is not hard to see that it can not be a submanifold of the torus.

While not all orbits turn out to be submanifolds, we just want to mention that the construction does not create new manifold structures on the orbits. One can establish the following theorem which we will not prove here. See [HN12, Corollary 10.1.16].

4.2.12 Proposition *Let $\sigma: G \times M \to M$ be a Lie group action and $m \in M$. Assume that for some $m \in M$ the orbit O_m is a submanifold of M, then $\overline{\sigma}^M : G/G_m \to O_m \subseteq M$ is a diffeomorphism.*

We know already that the canonical action of G on G/H is smooth and transitive. Now Proposition 4.2.12 shows that also the converse is true, if $\sigma: G \times M \to M$ is a transitive smooth action, then M is a homogeneous space (with the original manifold structure).

We end this section with three examples whihc are of independent interest both for their uses in geometry and since they appear in a wide variety of applications.

4.2.13 Example The orthogonal group $\mathrm{O}_n(\mathbb{R})$ acts smoothly on $\mathbb{R}^n$, it's orbits are the spheres (see e.g. Example 4.0.6)

$$S(r) := \{x \in \mathbb{R}^n : \|x\| = r\}, \quad r \geq 0.$$

All these spheres carry natural manifold structures. Therefore Proposition 4.2.12 implies that for each $r > 0$ we have

$$S(r) = \mathbb{S}^{n-1} = \mathrm{O}_n(\mathbb{R})/\mathrm{O}_n(\mathbb{R})_{e_1},$$

where

$$\mathrm{O}_n(\mathbb{R})_{e_1} = \left\{\begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix} : d \in \mathrm{O}_{n-1}(\mathbb{R})\right\} \cong \mathrm{O}_{n-1}(\mathbb{R}).$$

In particular, spheres in $\mathbb{R}^n$ are homogeneous spaces.

4.2.14 Example Consider $\mathbb{H} := \{z \in \mathbb{C} \mid \mathrm{Im} z > 0\}$ the upper half-plane in the complex numbers. Then we let $\mathrm{SL}_2(\mathbb{R})$ act via

$$\sigma: \mathrm{SL}_2(\mathbb{R}) \times \mathbb{H} \to \mathbb{H}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}.z := \frac{az+b}{cz+d}.$$

In Exercise 4.3 3. we show that this map yields a transitive Lie group action. Let us determine the stabiliser subgroup $\mathrm{SL}_2(\mathbb{R})_{\mathrm{i}}$ at the point $\mathrm{i} = \sqrt{-1}$. By definition $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is in this subgroup of $\mathrm{SL}_2(\mathbb{R})$ if and only if

$$\frac{a\mathrm{i}+b}{c\mathrm{i}+d} = \mathrm{i} \Leftrightarrow ad-bc = 1 \text{ and } a=d, b=-c.$$

In other words: The columns of the matrix are orthogonal to each other and of length 1, while the determinant is equal to 1. Therefore $\mathrm{SL}_2(\mathbb{R})_{\mathrm{i}} = \mathrm{SO}_2(\mathbb{R})$. We deduce that $\mathbb{H} = \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$ is a homogeneous space.

As a side-remark: The upper half-plane $\mathbb{H}$ and its homogeneous space structure are of interest for example as $\mathbb{H}$ is a model for hyperbolic geometry (which is a certain flavor of non-euclidean geometry).

4.2.15 Example (Graßmannians) Let $M := \mathrm{Gr}_k(\mathbb{R}^n)$ be the set of all k-dimensional linear subspaces of $\mathbb{R}^n$. We will construct a manifold structure on M turning it into a homogeneous space, and call M with this structure the *Graßmann manifold of degree k*.

Note first that we know from linear algebra that the action

$$\sigma\colon \mathrm{GL}_n(\mathbb{R}) \times \mathrm{Gr}_k(\mathbb{R}^n) \to \mathrm{Gr}_k(\mathbb{R}^n), \quad (g, F) \mapsto g(F)$$

is transitive. Let $F := \mathrm{span}\{e_1, \ldots, e_k\}$. Writing elements of $M_n(\mathbb{R})$ as 2×2 block matrices according to

$$M_n(\mathbb{R}) = \begin{pmatrix} M_k(\mathbb{R}) & M_{k,n-k}(\mathbb{R}) \\ M_{n-k,k}(\mathbb{R}) & M_{n-k}(\mathbb{R}) \end{pmatrix},$$

the stabiliser of F in $\mathrm{GL}_n(\mathbb{R})$ is

$$\mathrm{GL}_n(\mathbb{R})_F := \left\{\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : a \in \mathrm{GL}_k(\mathbb{R}), b \in M_{n,n-k}(\mathbb{R}), d \in \mathrm{GL}_{n-k}(\mathbb{R})\right\},$$

which is a closed subgroup. Then the homogeneous space $\mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{R})_F$ carries a natural manifold structure by Proposition 4.2.12. Since the orbit map of F induces a bijection

$$\overline{\sigma}^F\colon \mathrm{GL}_n(\mathbb{R})/\mathrm{GL}_n(\mathbb{R})_F \to \mathrm{Gr}_k(\mathbb{R}^n), \quad g\,\mathrm{GL}_n(\mathbb{R})_F \mapsto g(F),$$

we obtain a manifold structure on $\mathrm{Gr}_k(\mathbb{R}^n)$ by declaring this map a diffeomorphism. Note that the natural action of $\mathrm{GL}_n(\mathbb{R})$ then becomes smooth.

4.3 Exercises

1. A surjective map $f\colon X \to Y$ between topological spaces is called a ***quotient map*** if a subset $O \subseteq Y$ is open if and only if $f^{-1}(O)$ is open. Show that...

a) If X is a topological space and Y is a set which we endow with the final topology with respect to the map $f\colon X \to Y$, then f is a quotient map.

b) Let $f\colon X \to Y$ be a quotient map. Show that $g\colon Y \to Z$ (with Z a topological space) is continuous if and only if $g \circ f$ is continuous.

* **2.** ***(Whiteheads Lemma[3])*** Assume that $f\colon X \to Y$ is a quotient map and Z is a locally compact space, show that

$$\mathrm{id}_Z \times f\colon Z \times X \to Z \times Y, (z,x) \to (z, f(x))$$

is a quotient map.
Warning: In general the cartesian product of two quotient maps is not a quotient map. See [Mun00, Section 22] for examples and check out https://dantopology.wordpress.com/2023/04/21/the-product-of-the-identity-map-and-a-quotient-map/ for more information on how even established mathematicians fell into the trap of believing erroneously that this holds.
Hint: The standard proof uses local compactness and the implication (a) $\Rightarrow$ (b) of

Kuratowski's theorem (several results are called by this name)

The following statements are equivalent for Hausdorff topological spaces:

(a) The topological space K is compact

(b) for any topological space S the projection $K \times S \to S$ is a closed map.

3. Explain how Corollary 4.2.8 can be deduced from the proof of Theorem 4.2.6,

4. Let $\phi\colon G_1 \to G_2$ be a surjective Lie group morphism (e.g. the Heisenberg quotient, Example 4.2.9). Review the proof that $G_1/\ker\phi$ is a Lie group. When is the induced morphism $\phi\colon G_1/\ker\phi \to G_2$ an isomorphism of Lie groups?

5. Show that the action σ of $\mathrm{SL}_2(\mathbb{R})$ on the upper complex half-plane in Example 4.2.14 makes sense as a map and is transitive.

6. For a Lie group G we constructed in Exercise 3.5 1. the unit component G_0. Work out the quotient topology from Lemma 4.2.4 on the group of components $C := G/G_0$ and compare with the construction in Exercise 1.7 1. b).
Hint: Note that every connected component of G is of the form gG_0 for some $g \in G$.

[3] WARNING: This is a somewhat involved topological exercise. Only attempt this if you want a challenge and are confident of your skills in point set topology. The proof is recorded for example in [Eng89, Theorem 3.3.17].

A. Basics from earlier courses

In the following sections we collect basic material which we will use in this course. It should be familiar from earlier courses and we will not provide proofs.

A.1. Topology

An introductory text on general topology is [Mun00]. Recall that a *topological space* $(X, \mathcal{T})$ is a set X together with a system of subsets $\mathcal{T}$, such that

- $\emptyset, X \in \mathcal{T}$
- if $U_i \in \mathcal{T}$ for every $i \in I$ for some index set I, then $\bigcup_i U_i \in \mathcal{T}$
- if $U, V \in \mathcal{T}$, so is $U \cap V \in \mathcal{T}$

The system $\mathcal{T}$ is called the *topology* and elements $U \in \mathcal{T}$ are called *open sets*.

If $C = X \setminus U$ for $U \in \mathcal{T}$ (i.e. C is the complement of an open set), then C is a *closed set*. **Warning**: There are sets in topological spaces which are neither open nor closed.

A.1.1 Lemma *The following is true in any topological space* $(X, \mathcal{T})$*:*

- *the intersection of closed sets is closed.*
- *both $\emptyset$ and X are open and closed.*

We say that $U \overset{\circ}{\subseteq} X$ is an *open neighborhood* of x if $x \in U$. A (not necessarily open) set $N \subseteq X$ is called *neighborhood* of x if it contains an open neighborhood of x.

A.1.2 Example Recall that a metric space (X, d) is a set X together with a function $d\colon X \times X \to [0, \infty)$ which satisfies the following properties for all $x, y, z \in X$:

- $d(x, y) = 0$ if and only if $x = y$.
- $d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality)

Every metric space is in a canonical way a topological space. Its *metric topology* is obtained by declaring a set $U \subseteq X$ to be open if for every $x \in U$, there exists $r > 0$ such that the open ball

$$B_r(x) := \{y \in V \mid \|x - y\| < r\}$$

is contained in U.

A.1.3 Example For a normed space $(V, \|\cdot\|)$ define a metric via $d(x, y) := \|x - y\|$. Thus V is a topological spaces by Example A.1.2 and when we talk about the topology of a normed space (or the normed topology) we shall always use the associated metric topology.

Usually we care only for the normed space $\mathbb{R}^n$. In metric spaces we can test closedness via sequences:

A.1.4 Lemma *Let (X, d) a metric space. We say that a sequence $(x_n)_{n \in \mathbb{N}}$ converges to x in X if $\lim_{n \to \infty} d(x_n, x) = 0$.*

A subset $A \subseteq X$ is closed in X if and only if for every sequence $(x_n)_n \subseteq A$ which converges to x in X we must have $x \in A$.

A.1.5 Definition (Constructing new topological spaces from given ones) Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be topological spaces.

- If $S \subseteq X$ is a subset, it becomes a topological space with respect to the *subspace topology*. A set V is open in the subspace topology if there exists $U \in \mathcal{T}$ such that $V = U \cap S$.
- The cartesian product $X \times Y := \{(x, y) \mid x \in X, y \in Y\}$ is a topological space with respect to the so called *product topology*. A set $W \subseteq X \times Y$ is open if for every $(x, y) \in W$ there exist open neighborhoods $U_x \subseteq^{\circ} X$ of x and $V_y \subseteq^{\circ} Y$ of y, such that $U_x \times V_y \subseteq W$.
- Let Y be a set (without topology!) and $f_i \colon X_i \to Y, i \in I$ maps from topological spaes $(X_i, \mathcal{T}_i)$ to Y. We define a subset U of Y to be open if and only if $f_i^{-1}(U)$ is open in X_i for each $i \in I$. This gives rise to a topology called the *final topology (with respect to the f_i)*.

A.1.6 Definition Let $f \colon X \to Y$ be a map between topological spaces. Then f is *continuous* if for every $U \subseteq^{\circ} Y$ the preimage

$$f^{-1}(U) := \{x \in X \mid f(x) \in U\} \quad \text{is open in } X.$$

Note that if $f\colon (X,d) \to (Y,\tilde{d})$ is a map between metric spaces, then f is continuous in the sense of Definition A.1.6 if and only if it is continuous in the following sense

$$\text{for all } (x_n)_{n\in\mathbb{N}} \subseteq X \text{ with } x_n \to x,\ f(x_n) \to f(x) \text{ in } Y.$$

A.1.7 Definition A topological space X is called *compact*, if every open cover[1] $(U_i)_{i\in I}$ of X admits a finite subcover $(U_{i_j})_{j=1,2,\ldots,n}$.

Recall that a set $S \subseteq \mathbb{R}^n$ is bounded (by definition) if it is contained in a sufficiently large ball

$$B_r(0) := \{x \in \mathbb{R}^n : \|x\| < r\}.$$

In the earlier courses we learned that for subsets of $\mathbb{R}^n$ the following is equivalent to the above definition of compactness:

Let $X \subseteq \mathbb{R}^n$ for some $n \in \mathbb{N}$. Then X is compact if and only if

- every sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ has a convergent subsequence $(X_{n_k})_{k\in\mathbb{N}}$.
- X is closed and bounded (**Heine-Borel theorem**, [Mun00, Theorem 27.3]).

A.1.8 Example The closed balls $\overline{B_r(0)} = \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$ are compact.

Compact sets are stable under continuous maps, i.e. if $f\colon A \to B$ is a continuous map and $K \subseteq A$ is compact, then also $f(K)$ is compact. Also if C is a closed subset of a compact set K, then C is compact. (See e.g. [Mun00, Theorem 26.2 and Theorem 26.5])

A.1.9 Definition A subset U of a topological space X is called

- *path-connected* if for every pair of points $x, y \in U$ there exists a continuous path $\gamma\colon [0,1] \to U$ such that $\gamma(0) = x$ and $\gamma(1) = y$.
- *connected* if for every $C \subseteq U$ which is both closed and open, we have either $C = \emptyset$ or $C = U$.

If $x \in X$ we call the maximal set U_x which contains x and is path connected, the *path component of x* (in the topological space X). Similarly, the maximal connected subset of X which contains x is called the *connected compoent of x* in X.

The (path) components of the points partition a topological space into disjoint sets. Following general language we will call these disjoint sets the *(path) components* of the topological space.

[1] A system of subsets of a space X is called open cover if it consists of open subsets and the union of these subsets are already X.

Every path-connected subset of a topological space is connected. In general the converse is not true. However, for open subsets of $\mathbb{R}^n$ one can prove that path-connectedness and connectedness are equivalent, [Mun00, Theorem 25.5]. So in particular, this carries over to manifolds.

A.1 Exercises

1. Consider the reals $\mathbb{R}$ with their usual metric topology given by $d(x, y) = |x - y|$. Describe how open and closed subsets of $\mathbb{R}$ look in this topology. Find an example of a subset which is neither open nor closed.

2. Let $(E, \|\cdot\|)$ be a normed space endowed with the norm topology. Prove that open balls

$$B_r(x) = \{y \in E : \|x - y\| < r\}$$

are open subsets in this topology. Then prove that the normed topology is Hausdorff and in particular, the singletons $\{x\}$ are closed subsets.

3. Use the definition of continuity to prove that the composition $f \circ g : A \to C$ is continuous if $f : B \to C$ and $g : A \to B$ are continuous maps.

4. Let (X, T) be a topological space and $S \subseteq X$ a subset. Prove that the subspace topology...

a) turns S into a topological space.

b) turns the inclusion $i : S \to X$ into a continuous map and a mapping into $f : Y \to S$ is continuous if and only if $i \circ f : Y \to X$ is continuous.

5. Let $X \times Y$ be the cartesian product of topological spaces with the product topology.

a) Prove that the projections $\pi_X : X \times Y \to X, (x, y) \to x$ and $\pi_Y : X \times Y \to Y, (x, y) \to y$ are continuous.

b) Let $f = (f_X, f_Y) : A \to X \times Y$ be a mapping from a topological space into the product. Show that f is continuous if and only if the components $f_X : A \to X$ and $f_Y : A \to Y$ are continuous.

6. Let $f_i : (X_i, T_i) \to Y, i \in I$ be maps from topological spaces to a set Y. Endow Y with the final topology F with respect to the family $(f_i)_{i \in I}$. Prove that

a) The final topology as in Definition A.1.5 is indeed a topology.

b) the final topology makes all the mappings f_i continuous.

c) a map $g : (Y, F) \to (Z, T)$ is continuous if and only if $g \circ f_i$ is continuous for every $i \in I$.

7. Let X be a topological space which is path-connected. Prove that X is connected.

A.2. Calculus in several variables

We recall some basic results from calculus in several variables, see e.g. [HN12, Die69, Lan83].

A.2.1 Definition Let $n, m \in \mathbb{N}$ and $U \overset{\circ}{\subseteq} \mathbb{R}^n$ (recall that we denote open sets with this symbol). A function $f\colon U \to \mathbb{R}^m$ is called *differentiable* at $x \in U$ if there exists a linear map $L \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ such that for one norm (whence for all norms) on $\mathbb{R}^n$ we have

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - L(x)\|}{\|h\|} = 0 \tag{A.1}$$

If f is differentiable at x, then for each $h \in \mathbb{R}^n$ we have

$$\lim_{t=0} t^{-1}(f(x+th) - f(x)) = \lim_{t \to 0} t^{-1} L(th) = L(h)$$

so that $L(h)$ is the directional derivative of f at x in the direction h. Thus (A.1) determines L uniquely and we also write

$$\mathbf{d}f(x)(h) := \lim_{t \to 0} t^{-1}(f(x+th) - f(x)) = L(h)$$

and call the linear map $\mathbf{d}f(x)$ the *differential of f at x*.

If $e_1, \ldots, e_n$ is the standard basis of $\mathbb{R}^n$. then

$$\frac{\partial f}{\partial x_i}(x) := \mathbf{d}f(x)(e_i)$$

is called the *ith partial derivative of f at x*. If f is differentiable at each $x \in U$, then the partial derivatives are functions

$$\frac{\partial f}{\partial x_i}\colon U \to \mathbb{R}^m,$$

and we say that f is *continuously differentiable*, or a *C^1-map*, if all its partial derivativves are continuous. For $k \geq 2$, the map f is said to be a *C^k-map* if it is C^1 and all its partial derivatives are C^{k-1}-maps. We say that f is *smooth* or a *C^∞-map*, if it is C^k for each $k \in \mathbb{N}$. If f is only *continuous*, we also say f is a C^0-map.

Denote the space of C^k-maps $U \to \mathbb{R}^m$ by $C^k(U, \mathbb{R}^m)$ for $k \in \mathbb{N}_0$.

If $I \subseteq \mathbb{R}$ is an interval and $\gamma\colon I \to \mathbb{R}^n$ is a differentiable curve, we also write

$$\dot{\gamma}(t) = \gamma'(t) = \lim_{h \to 0} \frac{\gamma(t+h) - \gamma(t)}{h}.$$

Note that $\gamma'(t) = \mathbf{d}\gamma(t)(e_1)$, where $e_1 = 1 \in \mathbb{R}$ is the canonical basis of the reals.

A.2.2 Proposition (Rule on partial differentials) *Let $n_1, n_2, m \in \mathbb{N}$ and $U \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ and $f : U \to \mathbb{R}^m$ continuous. Then f is C^1 if and only if the limits*

$$\mathbf{d}_1 f(x,y)(v_1) := \lim_{t\to 0} t^{-1}(f(x+tv_1,y) - f(x,y)),$$
$$\mathbf{d}_2 f(x,y)(v_2) := \lim_{t\to 0} t^{-1}(f(x,y+tv_2) - f(x,y))$$

exists for all $(x,y) \in U$ and $(v_1,v_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ and yields continuous maps $\mathbf{d}_i f : U \to \mathrm{Lin}(\mathbb{R}^{n_i}, \mathbb{R}^m)$, $i = 1,2$. Moreover, one has

$$\mathbf{d}f(x,y)(v_1,v_2) = \mathbf{d}_1 f(x,y)(v_1) + \mathbf{d}_2 f(x,y)(v_2)$$

A.2.3 Definition (Higher derivatives) For $k \geq 2$, a C^k-map $f : U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$, higher derivatives are defined inductively y

$$\mathbf{d}^k f(x)(h_1, \ldots h_k) = \lim_{t\to 0} t^{-1}\left(\mathbf{d}^{k-1} f(x+th_k)(h_1,\ldots,h_{k-1}) - \mathbf{d}^{k-1} f(x)(h_1,\ldots,h_{k-1})\right).$$

We obtain continuous maps

$$\mathbf{d}^k f : U \times (\mathbb{R}^n)^k \to \mathbb{R}^m.$$

In terms of concrete coordinates and the standard basis $e_, \ldots e_n$ for $\mathbb{R}^n$ we then have

$$\mathbf{d}^k f(x)(e_{i_1}, \ldots, e_{i_k}) = \frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(x).$$

A.2.4 Theorem (Schwarz' Theorem) *Let $U \subseteq \mathbb{R}^m$ and if $f \in C^k(U, \mathbb{R}^n)$ and $k \geq 2$, then the functions $(h_1, \ldots, h_k) \to \mathbf{d}^k f(x)(h_1, \ldots, h_k)$, $x \in U$, are symmetric k-linear maps. In other words: the order of the h_i is irrelevant.*

A.2.5 Definition Let $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$. A map $f : U \to V$ is called C^k if it is C^k as a map $U \to \mathbb{R}^m$.

For $n \geq 1$, a C^k-map $f : U \to V$ is called a *C^k-diffeomorphism* if there exists a C^k-map $g : V \to U$ such that

$$f \circ g = \mathrm{id}_V \quad \text{and} \quad g \circ f = \mathrm{id}_U.$$

This is obviously equivalent to f being bijective and f^{-1} being a C^k-map. Whenever such a diffeomorphism exists, we say that the domains U and V are *C^k-diffeomorphic*. For $k = 0$, we obtain the notion of a *homeomorphism*.

A.2.6 Theorem (Chain rule) *Let $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$, $f : U \to V$ and $g : V \to \mathbb{R}$ be C^k-maps. Then $g \circ f$ is a C^k-map. and for each $x \in U$ we have (as linear maps):*

$$\mathbf{d}(g \circ f)(x) = \mathbf{d}g(f(x)) \circ \mathbf{d}f(x).$$

A.2.7 Theorem (Inverse function theorem) *Let $U \subseteq \mathbb{R}^n$, $x_0 \in U, k \in \mathbb{N} \cup \{\infty\}$, and $f: U \to \mathbb{R}^n$ a C^k-map for which the linear map $\mathrm{d}f(x_0)$ is invertible. Then there exists an open neighborhood V of x_0 in U for which $f|_V : V \to f(V)$ is a C^k-diffeomorphism onto an open subset of $\mathbb{R}^n$.*

A.2.8 Theorem (Implicit function theorem) *Let $U \subseteq \mathbb{R}^m \times \mathbb{R}^n$ and $g: U \to \mathbb{R}^m$ be a C^k-function, $k \in \mathbb{N} \cup \{\infty\}$. Further, let $(x_0, y_0) \in U$ with $g(x_0, y_0) = 0$ such that the linear map*

$$\mathrm{d}_1 g(x_0, y_0) \colon \mathbb{R}^m \to \mathbb{R}^m, \quad v \mapsto \mathrm{d}g(x_0, y_0)(v, 0)$$

is invertible. Then there exist open neighborhoods $V_1 \subseteq \mathbb{R}^m$ of x_0 and $V_2 \subseteq \mathbb{R}^n$ of y_0 with $V_1 \times V_2 \subseteq U$, and a C^k-function $f: V_2 \to V_1$ with $f(y_0) = x_0$ such that

$$\{(x, y) \in V_1 \times V_2 \colon g(x, y) = 0\} = \{(f(y), y) \colon y \in V_2\}.$$

For the last part we need vector valued integrals. For a function $f: [a, b] \to \mathbb{R}^n$ we write $f = (f_i)_{i=1,\ldots,n}$ for its components. Then we define

$$\int_a^b f(t)\,\mathrm{d}t := \begin{pmatrix} \int_a^b f_1(t)\,\mathrm{d}t \\ \vdots \\ \int_a^b f_n t)\,\mathrm{d}t \end{pmatrix}.$$

So we are actually only computing n real valued integrals (by looking at the components). In particular, we can thus compute the integral of a function along a curve using this definition. For basic integration theory consult [Die69, Lan83]. We need two results:

A.2.9 Lemma *(Continuous parameter dependence, [Die69, 8.11.1 and 8.7.8]) Let $a < b$ be real numbers and $U \subseteq \mathbb{R}^n$ such that $f: [a, b] \times U \to \mathbb{R}^m$ is a continuous function. Then the function*

$$F: U \to \mathbb{R}^m, \quad F(u) := \int_a^b f(t, u)\,\mathrm{d}t$$

is continuous. In particular, the following identity holds:

$$\lim_{u \to \tilde{u}} F(u) = \int_a^b \lim_{u \to \tilde{u}} f(t, u)\,\mathrm{d}t.$$

Proof. It suffices to prove this for the components, so without loss of generality $n = 1$. Since $[a, b]$ is compact, for any $u_0 \in U$ the functions $f(t, u_0) - f(t, u)$ is uniformly continuous in a neighborhood of u_0. Uniform continuity means that For every $\varepsilon > 0$ there is thus $\delta > 0$ such that $\|u_0 - u\| < \delta$ implies

$$|F(u) - F(u_0)| \leq \int_a^b |f(t, u_0) - f(t, u)|\,\mathrm{d}t \leq \sup_{t, \|u - u_0\| < \delta} |f(t, u) - f(t, u_0)|(b - a) < \varepsilon.$$

This establishes continuity and the formula for the limit. □

The next result is the $\mathbb{R}^n$ integral version of the familiar one-dimensional mean value theorem (for a proof apply the mean value theorem in every component).

A.2.10 Proposition *(Mean value theorem, integral form, [Lan83, 5, §4 Theorem 4.2]) Let $U \circ \mathbb{R}^n$ and $f: U \to \mathbb{R}^d$ be a C^1-map. Assume that for $x, y \in U$ also the line segment $\overline{xy} := \{tx + (1-t)y : t \in [0,1]\}$ is contained in U, then*

$$f(y) - f(x) = \int_0^1 \mathbf{d}f(x + t(y-x))(y-x)\,\mathrm{d}t \tag{A.2}$$

Note that in the statement of Proposition A.2.10 the vector $(y-x)$ is exactly the vector pointing from x to y, so we are really integrating $\mathbf{d}f$ along the line segment connecting x and y in the direction of the line segmnet. A standard identity we will need later on is given as Exercise below:

A.2 Exercises

1. Let $d, n, m \in \mathbb{N}$, $U \circ \mathbb{R}^n$, $V \circ \mathbb{R}^m$ and $f: U \to \mathbb{R}^d$ a C^2-map and both $g: V \to U$ and $h: V \to U$ be C^1. Prove that the differential of the C^1-map $\varphi := \mathbf{d}f \circ (g,h): V \to \mathbb{R}^d$, $\varphi(x) = \mathbf{d}f(g(x))(h(y))$ is given by

$$\mathbf{d}\varphi(x)(y) = \mathbf{d}^2 f(g(x))(h(x), \mathbf{d}g(x)(y)) + \mathbf{d}f(g(x))(\mathbf{d}h(x)(y)) \tag{A.3}$$

for $x \in V, y \in \mathbb{R}^m$.

A.3. Differential equations and vector fields

In this section we take a look at vector fields on open subsets of euclidean space and the differential equations they determine.

A.3.1 Remark Let $U \circ \mathbb{R}^n$ for some $n \in \mathbb{N}$. A function $F \in C^k(U, \mathbb{R}^n)$ is also called a *C^k-vector field* on U.

The idea is that a vector field pins to every point $x \in U$ a vector $F(x)$. Later, when dealing with vector fields on manifolds, we shall take a different point of view and consider functions of the following type

$$U \ni v \mapsto X(v) = (v, F(v)) \text{ with } F \in C^\infty(U, \mathbb{R}^n)$$

as vector fields. The smooth map F of such a vector field is then called *principal part* of the vector field X.

Obviously if we have a vector field, we can always deduce its principal part and conversely, if we have a smooth map F, we can construct the vector field $X(v) = (v, F(v))$. Thus on open subsets of $\mathbb{R}^n$ one tends to be sloppy and call both types of functions a vector field (it is usually clear from the context what is meant).

The point is that we can formulate differential equations usign vector fields:

A.3.2 Definition Let I be an open interval containing 0 and $\gamma\colon I \to U \subseteq \mathbb{R}^n$ be a differentiable map. We call γ and *integral curve* of a vector field $X \in C^k(U, \mathbb{R}^n)$ if

$$\gamma'(t) = X(\gamma(t)) \text{ for each } t \in I. \tag{A.4}$$

Note that (A.4) implies that γ' is continuous and if γ is C^ℓ for any $\ell \leq k$, then also γ' is C^ℓ. Therefore, integral curves of a C^k-vector field are automatically C^{k+1}-maps.

If $J \supseteq I$ is an interval containing I, then an integral curve $\eta\colon J \to U$ is called an *extension* og γ if $\eta|_I = \gamma$. An integral curve γ is said to be *maximal* if it has no proper extensions.

A.3.3 Remark A curve $\gamma\colon I \to U$ is an integral curve of X if and only if $\tilde{\gamma}(t) \coloneqq \gamma(-t)$ is an integral curve of the vector field $-X$. More generally, for $a, b \in \mathbb{R}$, the curve $\eta(t) \coloneqq \gamma(at + b)$ is an integral curve of the vector field aX.

A.3.4 Definition For a continuous curve $\gamma\colon (a, b) \to U \subseteq \mathbb{R}^n$ we say that $\lim_{t\to b} \gamma(t) = \infty$ (γ "escapes to ∞") if for each compact $K \subseteq U$ there is $c \in (a, b)$ such that for all $t > c$ we have $\gamma(t) \notin K$. Similarly we write $\lim_{t\to a} \gamma(t) = \infty$ if for each $K \subseteq U$ compact there is $c \in (a, b)$ with $\gamma(t) \notin K$ for all $t < c$.

Recall the following result from another course on ODEs (or see [Ama90, Section 2]):

A.3.5 Theorem (Existence and Uniqueness of Integral curves) *Let $X \in C^k(U, \mathbb{R}^n)$, $k \geq 1$ for $U \subseteq \mathbb{R}^n$ and $p \in U$. Then there exists a unique maximal integral curve $\gamma_p\colon I_p \to U$ with $\gamma_p(0) = p$. If $a \coloneqq \inf I_p > -\infty$, then $\lim_{t\to a} \gamma_p(t) = \infty$ and if $b \coloneqq \sup I_p < \infty$, then $\lim_{t\to b} \gamma_p(t) = \infty$.*

A.3.6 Example Let us consider $U = \mathbb{R}$ and the vector field $X(s) = 1 + s^2$. The corresponding ordinary differential equation is

$$\gamma'(t) = X(\gamma(t)) = 1 + (\gamma(s))^2.$$

For the initial condition $\gamma(0) = 0$, the function $\gamma(t) = \tan(t)$ on $I =]\frac{-\pi}{2}, \frac{\pi}{2}[$ is the unique maximal solution and we note that $\lim_{t\to\pm\pi/2}\gamma(t) = \pm\infty$.

A.3.7 Example Let $U = (-1, 1)$ be the open intervall and $X(t) := 1$. Then $\gamma'(t) = X(\gamma(t)) = 1$ has a unique maximal solution $\gamma(t) = t$ on $(-1, 1)$. Note that also here $\lim_{t\to\pm1}\gamma(t) = \infty$ (even though the solution does not get infinitely large). If we instead would have considered the open set $U = \mathbb{R}$, then a maximal solution exists on all of $\mathbb{R}$.

A.3.8 Definition A vector field $C^\infty(U, \mathbb{R}^n)$ is called *complete* if for every $p \in U$ the maximal integral curve γ_p is defined on $I_p = \mathbb{R}$.

Local flows of vector fields

A.3.9 Definition Let $U \circ \mathbb{R}^n$. A *local flow* on U is a smooth map $\Phi\colon \Omega \to U$, where $\Omega \circ \mathbb{R} \times U$ is a set containing $\{0\} \times U$, such that for each $x \in U$ the intersection $I_x := \Omega \cap (\mathbb{R} \times \{x\})$ is an interval containing 0 and

$$\Phi(0, x) = x \text{ and } \Phi(t, \Phi(s, x)) = \Phi(t + s, x)$$

hold for all t, s, x for which both sides of the equation are defined. The maps $\alpha_x\colon I_x \to U, t \to \Phi(t, x)$ are called *flow lines*. The flow Φ is said to be *global* if $\Omega = \mathbb{R} \times U$.

A.3.10 Lemma *If $\Phi\colon \Omega \to U$ is a local flow, then*

$$X^\Phi(x) := \frac{d}{dt}\Big|_{t=0} \Phi(t, x) = \alpha_x'(0)$$

defines a smooth vector field. It is called the velocity field *or* infinitesimal generator *of the local flow Φ.*

A.3.11 Lemma *If $\Phi\colon \Omega \to U$ is a local flow on U, then the flow lines are integral curves of the vector field X^Φ. In particular, the local flow Φ is uniquely determined by the vector field X^Φ.*

Proof. If $\alpha_x\colon I_x \to U$ is a flow line and $s \in I_x$. Then for sufficiently small $t \in \mathbb{R}$,

$$\alpha_x(s + t) = \Phi(s + t, x) = \Phi(t, \Phi(s, x)) = \Phi(t, \alpha_x(s)).$$

Derivating in $t = 0$ we obtain $\alpha_x'(s) = X^\Phi(\alpha_x(s))$. That the flow is uniquely determined by X^Φ follows from Theorem A.3.5. □

A.3.12 Theorem *Each $X \in C^\infty(U, \mathbb{R}^n)$ is the velocity field of a unique local flow Fl^X defined by*

$$\mathrm{Fl}\colon D_X := \bigcup_{x \in U} I_x \times \{x\} \to U, \quad \mathrm{Fl}^X(t,x) := \gamma_x(t) \text{ for } (t,x) \in D_X,$$

where $\gamma_x \colon I_x \to U$ is the unique maximal integral curve of X with $\gamma_x(0) = x \in U$.

The proof of Theorem A.3.12 is a bit technical and can be found together with the extension we record in Proposition A.3.13 below in [HN12, Theorem 8.5.12 and Proposition 8.5.5]. We indeed need a slightly more advanced version in which the vector fields and their flows may depend on an additional parameter:

A.3.13 Proposition *Let $U \overset{\circ}{\subseteq} \mathbb{R}^n$ and $P \overset{\circ}{\subseteq} \mathbb{R}^k$. Assume that $\Psi \colon P \to C^\infty(U, \mathbb{R}^n)$ is a map such that the map*

$$\hat{\Psi}\colon P \times U \to \mathbb{R}^n, (p,x) \to \Psi(p)(x)$$

is smooth (the vector field $\Psi(p)$ depends smoothly on p). Then there exists for each $(p_0, x_0) \in P \times U$ an open neighborhood W of p_0 in P, an open interval $I \subseteq \mathbb{R}$ containing 0, an open neighborhood V of x_0 in U and a smooth map

$$\Phi\colon I \times W \times V \to U$$

such that for each $(p,x) \in W \times V$ the curve

$$\Phi^p_x\colon I \to U, \quad t \to \Phi(t,p,x)$$

is an integral curve of the vector field $\Psi(p)$ with $\Phi^p_x(0) = x$.

A.3.14 Corollary *In the situation of Proposition A.3.13 assume that all the vector fields $\Psi(p)$ are complete, then there exists a smooth map*

$$\Phi\colon \mathbb{R} \times P \times U \to U$$

such that for each $(p,x) \in P \times U$ the curve $\Phi^p_x\colon \mathbb{R} \to U, t \to \Phi(t,p,x)$ is an integral curve of the vector field $\Psi(p)$ with $\Phi^p_x(0) = x$.

B. Topological groups

In this appendix we collect material on topological groups. While we will study a more restricted class of examples. Every Lie group is a topological group but there are topological groups which are not Lie groups. Indeed the question of which topological groups are Lie groups is known as Hilbert's fifth problem[1].

Recall from Definition 1.1.3 that a topological group is a group which is also a topological space such that the group operations are continuous. Further we note that

$$\delta\colon G \times G \to G, \delta(g,h) := gh^{-1} \tag{B.1}$$

is continuous as a composition of the continuous group operations.

B.0.1 Lemma *Let G be a topological group and H a subgroup. Then the subspace topology turns H into a topological group.*

Proof. The subspace topology makes the inclusion $i\colon H \to G$ continuous and the group multiplication $m_H\colon H \times H \to H$ is continuous if $i \circ m_H$ is continuous by Exercise A.1.2. However, the group multiplication $m_G\colon G \times G \to G$ is continuous. Then

$$i \circ m_H = m_G \circ (i \times i)$$

is continuous. Similarly, the inversion ι_G is continuous and $i \circ \iota_H = \iota_G \circ i$ is continuous. We deduce that H is a topological group. □

B.0.2 Lemma *Let G be a topological group and $\mathbf{1}\quad U \circ G$. Then there exists $1\quad V \circ G$ such that $V \cdot V \quad U$ and $V = V^{-1}$ (we say V is* symmetric*).*

Proof. Multiplication $m\colon G \times G \to G$ and inversion $\iota\colon G \to G$ are continuous. By continuity if $\mathbf{1}\quad U \circ G$, there are $V_1 \times V_2 \circ G \times G$ with $m(V_1 \times V_2) \quad U$. Set $V := V_1 \cap V_2 \cap \iota(V_1 \cap V_2)$, then $\mathbf{1}\quad V \circ G$ and $V = \iota(V)$ such that $V \cdot V := m(V \times V) \quad U$. □

B.0.3 Lemma *Let G be a topological group such that the topology on G is Hausdor . Then for every $g \quad G$ and $C \quad G$ closed such that $g \quad C$ there exist open sets $g \quad U$ and $C \quad V$ such that $U \cap V = \ .$*

[1] See [Tao14] for a book account of the problem, its history and solution.

Proof. Since C is closed $O = G \setminus C$ is open. Now $g \in O$ whence $(g, \mathbf{1}) \in \delta^{-1}(O) \subseteq G \times G$. As δ is continuous, $\delta^{-1}(O)$ is open in $G \times G$. By definition of the product topology, we can thus find an open g-neighborhood U and an open $\mathbf{1}$-neighborhood W such that $U \times W \subseteq \delta^{-1}(O)$.

Consider now $y \in U \cap CW$, then $y = cw$ for some $w \in W, c \in C$. Now $(y, w) \in U \times W \subseteq \delta^{-1}(O)$, so $G \setminus C = O \ni y \cdot w^{-1} = cww^{-1} = c \in C$. This is a contradiction, whence $U \cap CW = \emptyset$. By Exercise B.1 2. CW is an open set which contains C (as W is a $\mathbf{1}$-neighborhood) and U is an open set containing g. □

B.1 Exercises

1. Let G be a topological group and $g \in G$.

a) Show that the left translation $\lambda_g : G \to G, h \mapsto g \cdot h$ is a homeomorphism (i.e. a bijective continuous map whose inverse is also continuous).

b) Is the same true for the right translation $\rho_g : G \to G, h \mapsto h \cdot g$?

c) Prove that inversion $\iota : G \to G, g \mapsto g^{-1}$ is a homeomorphism.

2. Let G be a topological group, with a Hausdorff topology. If $O \subset^{\circ} G$ and $S \subseteq G$ some subset. Show that the product sets

$$OS := \{o \cdot s : o \in O, s \in S\} \text{ and } SO := \{s \cdot o : o \in O, s \in S\}$$

are open in G.

3. Let G be a group which is also a topological space, such that the group product is continuous. Assume that there is $x \in G$ and an open x-neighborhood U_x such that inversion is continuous on U_x. Prove that G is a topological group.

4. Recall that for groups G_i, $1 \leq i \leq n$, the product set $G = G_1 \times G_2 \times \cdots \times G_n$ has a natural group structure given by

$$(g_1, \ldots, g_n)(h_1, \ldots, h_n) = (g_1 h_1, \ldots, g_n h_n).$$

The group G is called the ***direct product*** of the groups G_j. Show that if every G_j is a topological group, then the direct product G is a topological group with respect to the product topology.

C. The Heisenberg quotient is not linear

In this appendix we will prove the following:

C.0.1 Proposition *The Heisenberg quotient*

$$Q = \mathbb{R}^2 \times \mathbb{S}^1, \quad (x,y,z)\cdot(a,b,c) = (x+a, y+b, z\cdot c\cdot e^{\mathrm{i}xb})$$

from Example 3.0.8 is not isomorphic to a subgroup of $\mathrm{GL}_n(\mathbb{K})$ *for any* $n \in \mathbb{N}$.

In essence the argument will boil down to linear algebra. See e.g. [HN12, Example 9.5.20] for an alternative argument which uses more Lie theory (but also uses the Jordan decomposition we need to invoke in the following proof). We need two auxiliary lemmata.

C.0.2 Lemma *Let* $X \in M_n(\mathbb{C})$ *and* $p \in \mathbb{N}$ *such that* $X^p = I_n$. *Then* X *is diagonalisable over* $\mathbb{C}$ *and all eigenvalues* λ *of* X *are pth roots of the unit, i.e.* $\lambda^p = 1$.

Proof. If $\lambda \in \mathbb{C}$ is an eigenvalue of X with eigenvector $v \in \mathbb{C}^n \setminus \{0\}$. Then

$$v = I_n v = X^p v = X^{p-1}(Xv) = \lambda X^{p-1} v = \cdots = \lambda^p v.$$

We deduce that $\lambda^p = 1$.

To see that X is diagonalisable, we use dark magic and invoke the Jordan normal form theorem[1]. This means that we can find complex linear subspaces $\mathbb{C}^n = E_1 \times \cdots E_i, i \in \mathbb{N}$ which are invariant under X (i.e. X maps them to themselves) and on each E_j the matrix is of the form $X|_{E_j} = \lambda_j I_j + N_j$, where I_j is the identity matrix of size $m = \dim E_j$ and N_j is an upper triangular matrix. So this means that X is diagonalisable if we can prove that all the N_j are zero. Since I_j and N_j commute we apply the binomial theorem to

$$I_j = (X|_{E_j})^p = (\lambda_j I_j + N_j)^p = \binom{p}{0}\lambda_j^p I_j + \sum_{k=1}^{p}\binom{p}{k}\lambda_j^{p-k}N_j^k$$

Note that the first summand is equal to I_j. If $N_j = 0$, then the matrices $N_j, N_j^2, \ldots, N_j^m$ are all linearly independent, whence $\binom{p}{1}\lambda^{p-1}N_j = 0$. So this only happens if $N_j = 0$ (as λ_j is a pth root of the unit!). We deduce that X must be diagonalisable. □

[1] We will NOT prove the Jordan normal form theorem which is taught in TMA4145 - Linear Methods (or take your favorite linear algebra book or [Mey23, 7.8]).

C.0.3 Lemma *Let n N and p a prime number. Assume that there are A, B, C $\mathrm{GL}_n(\mathrm{C})$ with the following properties:*

- $ABA^{-1}B^{-1} = C$
- $[A, C] = 0, \quad [B, C] = 0.$
- $C^p = I_n$ *and p is the smallest number with this property.*

Then $n \geq p$.

Proof. By Lemma C.0.2, C is diagonalisable, hence we may choose a basis for C^n such that C actually is a diagonal matrix. If every eigenvalue λ of C equals 1, we would have $C = I_n$ but this contradicts that $C^p = I_n$ and $p > 1$ is the smallest number with this property. We may thus pick an eigenvalue $\lambda = 1$ with $\lambda^p = 1$. Since p is prime (and thus only is divided by 1 and p), and $\lambda = 1$ we deduce that p must be the smallest number such that $\lambda^p = 1$.

Let $E_\lambda = \{v \quad \mathrm{C}^n \mid Cv = \lambda v\}$ be the λ-eigenspace for C. We claim that $A(E_\lambda) \quad E_\lambda$ and $B(E_\lambda) \quad E_\lambda$. To see this, let $w \quad E_\lambda$, then since $[A, C] = 0$ we have

$$C(Aw) = A(Cw) = A(\lambda w) = \lambda(Aw).$$

In other words $Aw \quad E_\lambda$. The same argument works for B. We can thus restrict the matrices to linear maps on the eigenspace E_λ and note that $C|_{E_\lambda} = \lambda I_k$ and further

$$\lambda I_k = C|_{E_\lambda} = (ABA^{-1}B^{-1})|_{E_\lambda} = A|_{E_\lambda} B|_{E_\lambda} A^{-1}|_{E_\lambda} B|_{E_\lambda}.$$

Applying the determinant to both sides we obtain from the usual rules for the determinant

$$\det(A|_{E_\lambda}) \cdot \det(B|_{E_\lambda}) \cdot \frac{1}{\det(A|_{E_\lambda})} \cdot \frac{1}{\det(B|_{E_\lambda})} = \det(\lambda I_k) = \lambda^k.$$

However, the left hand side equals 1, whence $\lambda^k = 1$. As p is the smallest number with this property, we see that $k \geq p$ and since E_λ is a subspace of C^n we must have $n \geq p$. □

Proof of Proposition C.0.1. Let p N any prime number. Consider the elements $a := (1, 0, 1)$, $b := (0, \frac{2\pi}{p}, 1)$ and $c := a \cdot b \cdot a^{-1} \cdot b^{-1}$ in Q. A computation shows that

$$c = a \cdot b \cdot a^{-1} \cdot b^{-1} = (0, 0, e^{\frac{2\pi i}{p}}).$$

Clearly $c^p = (0, 0, 1)$ (the unit of Q!) and p is the smallest number with this property for our choice of c. Another easy computation shows that c commutes with all other elements in Q, whence in particular with a, b.

Assume now that $\alpha \colon Q \to \mathrm{GL}_n(\mathrm{C})$ is an injective group homomorphism. Then $A := \alpha(a), B := \alpha(b), C := \alpha(c)$ are matrices in $\mathrm{GL}_n(\mathrm{C})$ which satisfy the assumptions of Lemma C.0.3. We deduce that $n \geq p$. However, p was an arbitrary prime number and this contradicts the existence of n N. □

Bibliography

[Ama90] Amann, H. *Ordinary di erential equations. An introduction to nonlinear analysis.*, De Gruyter Stud. Math., vol. 13 (Berlin etc.: Walter de Gruyter, 1990)

[Die69] Dieudonné, J. *Foundations of modern analysis. Enlarged and corrected printing*. New York-London: Academic Press. xv, 387 p. (1969). 1969

[DK00] Duistermaat, J. J. and Kolk, J. A. C. *Lie groups*. Universitext (Berlin: Springer, 2000)

[Eng89] Engelking, R. *General topology.*, Sigma Ser. Pure Math., vol. 6 (Berlin: Heldermann Verlag, 1989), rev. and compl. ed. edn.

[Hal15] Hall, B. *Lie groups, Lie algebras, and representations. An elementary introduction*, Grad. Texts Math., vol. 222 (Cham: Springer, 2015), 2nd ed. edn.

[Hel01] Helgason, S. *Di erential geometry, Lie groups, and symmetric spaces.*, Grad. Stud. Math., vol. 34 (Providence, RI: AMS, 2001)

[HN12] Hilgert, J. and Neeb, K.-H. *Structure and geometry of Lie groups*. Springer Monogr. Math. (Berlin: Springer, 2012). doi:10.1007/978-0-387-84794-8

[Hum80] Humphreys, J. E. *Introduction to Lie algebras and representation theory. 3rd printing, rev*, Grad. Texts Math., vol. 9 (Springer, Cham, 1980)

[Hyd00] Hydon, P. E. *Symmetry methods for di erential equations. A beginner's guide*. Camb. Texts Appl. Math. (Cambridge: Cambridge University Press, 2000)

[Jac85] Jacobson, N. *Basic algebra I. 2nd ed*. New York: W. H. Freeman and Company. XVIII, 499 p. £ 19.95 (1985). 1985

[Kna02] Knapp, A. W. *Lie groups beyond an introduction*, Prog. Math., vol. 140 (Boston, MA: Birkhäuser, 2002), 2nd ed. edn.

[Lan83] Lang, S. *Real analysis. 2nd ed*. Reading, Massachusetts, etc.: Addison-Wesley Publishing Company, Advanced Book Program/World Science Division. XIV, 533 p. $ 23.95 (1983). 1983

[Lee13] Lee, J. M. *Introduction to smooth manifolds*, Grad. Texts Math., vol. 218 (New York, NY: Springer, 2013), 2nd revised ed edn. doi:10.1007/978-1-4419-9982-5. URL zenodo.org/record/4461500

[Mey23] Meyer, C. D. *Matrix analysis and applied linear algebra*, Other Titles Appl. Math., vol. 188 (Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2023), 2nd edition edn. doi:10.1137/1.9781611977448

[Mun00] Munkres, J. R. *Topology.* (Prentice Hall, 2000), 2. edn.

[Olv93] Olver, P. J. *Applications of Lie groups to di erential equations.*, Grad. Texts Math., vol. 107 (New York: Springer-Verlag, 1993), 2nd ed. edn.

[OM01] Owren, B. and Marthinsen, A. *Integration methods based on canonical coordinates of the second kind*. Numer. Math. **87** (2001)(4):763–790

[Sch23] Schmeding, A. *An introduction to infinite-dimensional di erential geometry*, Camb. Stud. Adv. Math., vol. 202 (Cambridge: CUP, 2023)

[Sta07] Starrett, J. *Solving di erential equations by symmetry groups*. Am. Math. Mon. **114** (2007)(9):778–792. doi:10.1080/00029890.2007.11920470

[Tao14] Tao, T. *Hilbert's fifth problem and related topics*, Grad. Stud. Math., vol. 153 (Providence, RI: American Mathematical Society (AMS), 2014)