Lie Groups

Karl-Hermann Neeb

October 12, 2010

Contents

Notation

$\mathbb{N} := \{1, 2, 3, ...\}$ natural numbers
$\mathbb{K}^\times := \{x \in \mathbb{K} \colon x \neq 0\}$, $\mathbb{K}$ a field
$\mathcal{A}^\times := \{x \in \mathcal{A} \colon (\exists y \in R) xy = yx = \mathbf{1}\}$, unit group of a unital algebra

For subsets $A, B \subseteq G$ of a group:
$A^{-1} := \{a^{-1} \colon a \in A\}$
$AB := \{ab \colon a \in A, b \in B\}$
The identity element of a group G is usually denoted $\mathbf{1}$. If G is abelian and the product is written as addition, we write 0 for the identity element.
For $A = (a_{ij})_{i,j=1,\ldots,n} \in M_n(\mathbb{C})$: $A^\top = (a_{ji})$, $\overline{A} = (\overline{a_{ij}})$, $A^* = \overline{A}^\top = (\overline{a_{ji}})$.

Chapter 1

Topological Groups

1.1 Definitions and Examples

Definition 1.1.1. A *topological group* is a pair (G, τ) of a group G and a Hausdorff topology τ for which the group operations

$$m_G \colon G \times G \to G, \quad (x, y) \mapsto xy \quad \text{and} \quad \eta_G \colon G \to G, \quad x \mapsto x^{-1}$$

are continuous if $G \times G$ carries the product topology. Then we call τ a *group topology* on the group G.

Remark 1.1.2. The continuity of the group operations can also be translated into the following conditions which are more direct than referring to the product topology on G. The continuity of the multiplication m_G in $(x, y) \in G \times G$ means that for each neighborhood V of xy there exist neighborhoods U_x of x and U_y of y with $U_x U_y \subseteq V$. Similarly, the continuity of the inversion map η_G in x means that for each neighborhood V of x^{-1}, there exist neighborhoods U_x of x with $U_x^{-1} \subseteq V$.

Remark 1.1.3. For a group G with a topology τ, the continuity of m_G and η_G already follows from the continuity of the single map

$$\varphi \colon G \times G \to G, \quad (g, h) \mapsto gh^{-1}.$$

In fact, if φ is continuous, then the inversion $\eta_G(g) = g^{-1} = \varphi(\mathbf{1}, g)$ is the composition of φ and the continuous map $G \to G \times G, g \mapsto (\mathbf{1}, g)$. The continuity of η_G further implies that the product map

$$\mathrm{id}_G \times \eta_G \colon G \times G \to G \times G, \quad (g, h) \mapsto (g, h^{-1})$$

is continuous, and therefore $m_G = \varphi \circ (\mathrm{id}_G \times \eta_G)$ is continuous.

Remark 1.1.4. Every subgroup of a topological group is a topological group with respect to the subspace topology.

Example 1.1.5. (1) The additive group $(X,+)$ of every normed space $(X, \|\cdot\|)$ is a topological group because addition and negation are continuous maps. In particular, $(\mathbb{R}^n,+)$ is an abelian topological group with respect to any metric defined by a norm.

(2) $(\mathbb{C}^\times, \cdot)$ is a topological group and the circle group $\mathbb{T} := \{z \in \mathbb{C}^\times : |z| = 1\}$ is a compact subgroup.

(3) The group $\mathrm{GL}_n(\mathbb{R})$ of invertible $(n \times n)$-matrices is a topological group with respect to matrix multiplication. The continuity of the inversion follows from Cramer's Rule, which provides an explicit formula for the inverse in terms of determinants: For $g \in \mathrm{GL}_n(\mathbb{R})$, the inverse of g is given by

$$(g^{-1})_{ij} = \frac{(-1)^{i+j}}{\det g} \det(g_{mk})_{m \neq j, k \neq i}.$$

(see Proposition 1.1.9 for a different argument).

(4) Any group G is a topological group with respect to the discrete topology.

Lemma 1.1.6. *Let G be a topological group. Then the following assertions hold:*

(i) *The left multiplication maps $\lambda_g \colon G \to G, x \mapsto gx$ are homeomorphisms.*

(ii) *The right multiplication maps $\rho_g \colon G \to G, x \mapsto xg$ are homeomorphisms.*

(iii) *The conjugation maps $c_g \colon G \to G, x \mapsto gxg^{-1}$ are homeomorphisms.*

(iv) *The inversion map $\eta_G \colon G \to G, x \mapsto x^{-1}$ is a homeomorphism.*

Proof. (i) The continuity of the multiplication map implies by restriction that the maps λ_g are continuous. Since $\lambda_{g^{-1}}$ is the inverse of λ_g, it follows that each λ_g is a topological isomorphism, i.e., a homeomorphism.

(ii) is proved as in (i).

(iii) follows from (i) and (ii).

(iv) follows from the continuity of η_G and $\eta_G^2 = \mathrm{id}_G$. □

We have already argued above that the group $\mathrm{GL}_n(\mathbb{R})$ carries a natural group topology. This group is the unit group of the algebra $M_n(\mathbb{R})$ of real $(n \times n)$-matrices. As we shall see now, there is a vast generalization of this construction.

Definition 1.1.7. A *Banach algebra* is a triple $(\mathcal{A}, m_\mathcal{A}, \|\cdot\|)$ of a Banach space $(\mathcal{A}, \|\cdot\|)$, together with an associative bilinear multiplication

$$m_\mathcal{A} \colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}, (a,b) \mapsto ab$$

for which the norm $\|\cdot\|$ is *submultiplicative*, i.e.,

$$\|ab\| \leq \|a\| \cdot \|b\| \quad \text{for} \quad a, b \in \mathcal{A}.$$

By abuse of notation, we shall call $\mathcal{A}$ a Banach algebra, if the norm and the multiplication are clear from the context.

A *unital Banach algebra* is a pair $(\mathcal{A}, \mathbf{1})$ of a Banach algebra $\mathcal{A}$ and an element $\mathbf{1} \in \mathcal{A}$ satisfying $\mathbf{1}a = a\mathbf{1} = a$ for each $a \in \mathcal{A}$.

The subset

$$\mathcal{A}^\times := \{a \in \mathcal{A} \colon (\exists b \in \mathcal{A})\ ab = ba = \mathbf{1}\}$$

is called the *unit group of* $\mathcal{A}$ (cf. Exercise 1.1.8).

Remark 1.1.8. In a Banach algebra $\mathcal{A}$, the multiplication is continuous because $a_n \to a$ and $b_n \to b$ implies $\|b_n\| \to \|b\|$ and therefore

$$\|a_n b_n - ab\| = \|a_n b_n - ab_n + ab_n - ab\| \leq \|a_n - a\| \cdot \|b_n\| + \|a\| \cdot \|b_n - b\| \to 0.$$

In particular, left and right multiplications

$$\lambda_a \colon \mathcal{A} \to \mathcal{A}, x \mapsto ax, \quad \text{and} \quad \rho_a \colon \mathcal{A} \to \mathcal{A}, x \mapsto xa,$$

are continuous with

$$\|\lambda_a\| \leq \|a\| \quad \text{and} \quad \|\rho_a\| \leq \|a\|.$$

Proposition 1.1.9. *The unit group $\mathcal{A}^\times$ of a unital Banach algebra is an open subset and a topological group with respect to the topology defined by the metric $d(a,b) := \|a - b\|$.*

Proof. The proof is based on the convergence of the Neumann series $\sum_{n=0}^\infty x^n$ for $\|x\| < 1$. For any such x we have

$$(\mathbf{1} - x) \sum_{n=0}^\infty x^n = \Big(\sum_{n=0}^\infty x^n\Big)(\mathbf{1} - x) = \mathbf{1},$$

so that $\mathbf{1} - x \in \mathcal{A}^\times$. We conclude that the open unit ball $B_1(\mathbf{1})$ is contained in $\mathcal{A}^\times$.

Next we note that left multiplications $\lambda_g \colon \mathcal{A} \to \mathcal{A}$ with elements $g \in \mathcal{A}^\times$ are continuous (Remark 1.1.8), hence homeomorphisms because $\lambda_g^{-1} = \lambda_{g^{-1}}$ is also continuous. Therefore $gB_1(\mathbf{1}) = \lambda_g B_1(\mathbf{1}) \subseteq \mathcal{A}^\times$ is an open subset, showing that g is an interior point of $\mathcal{A}^\times$. Therefore $\mathcal{A}^\times$ is open.

The continuity of the multiplication of $\mathcal{A}^\times$ follows from the continuity of the multiplication on $\mathcal{A}$ by restriction and corestriction (Remark 1.1.8). The continuity of the inversion in $\mathbf{1}$ follows from the estimate

$$\|(\mathbf{1} - x)^{-1} - \mathbf{1}\| = \|\sum_{n=1}^\infty x^n\| \leq \sum_{n=1}^\infty \|x\|^n = \frac{1}{1 - \|x\|} - 1 = \frac{\|x\|}{1 - \|x\|},$$

which tends to 0 for $x \to 0$. The continuity of the inversion in $g_0 \in \mathcal{A}^\times$ now follows from the continuity in $\mathbf{1}$ via

$$g_n^{-1} - g^{-1} = g^{-1}(gg_n^{-1} - \mathbf{1}) = g^{-1}((g_n g^{-1})^{-1} - \mathbf{1})$$

because left and right multiplication with g^{-1} is continuous. This shows that $\mathcal{A}^\times$ is a topological group (cf. Exercise 1.1.2 for shortcuts in this argument). □

Example 1.1.10. (a) If $(X, \|\cdot\|)$ is a Banach space, then the space $\mathcal{L}(X)$ of continuous linear operators $A\colon X \to X$ is a unital Banach algebra with respect to the *operator norm*

$$\|A\| := \sup\{\|Ax\| \colon x \in X, \|x\| \leq 1\}$$

and composition of maps. Note that the submultiplicativity of the operator norm, i.e.,

$$\|AB\| \leq \|A\| \cdot \|B\|,$$

is an immediate consequence of the estimate

$$\|ABx\| \leq \|A\| \cdot \|Bx\| \leq \|A\| \cdot \|B\| \cdot \|x\| \quad \text{for} \quad x \in X.$$

In this case the unit group is also denoted $\mathrm{GL}(X) := \mathcal{L}(X)^\times$.

(b) Specializing (a) to $\mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$), we see that the algebra $\mathcal{A} = M_n(\mathbb{K})$ of $(n \times n)$-matrices with entries in $\mathbb{K}$ is a Banach algebra with respect to the operator norm

$$\|A\| := \sup\{\|Ax\| \colon \|x\| \leq 1, x \in \mathbb{K}^n\|,$$

where $\|\cdot\|$ is any norm on $\mathbb{K}^n$.

(c) If X is a compact space and $\mathcal{A}$ a Banach algebra, then the space $C(X, \mathcal{A})$ of $\mathcal{A}$-valued continuous functions on X is a Banach algebra with respect to pointwise multiplication $(fg)(x) := f(x)g(x)$ and the norm $\|f\| := \sup_{x \in X} \|f(x)\|$ (Exercise 1.1.7). Its unit group is

$$C(X, \mathcal{A})^\times = C(X, \mathcal{A}^\times),$$

because the continuity of the inversion in $\mathcal{A}^\times$ implies that for each $\mathcal{A}^\times$-valued function f, the pointwise inverse also is continuous.

(d) An important special case of (b) arises for $\mathcal{A} = M_n(\mathbb{C})$, where we obtain $C(X, M_n(\mathbb{C}))^\times = C(X, \mathrm{GL}_n(\mathbb{C})) = \mathrm{GL}_n(C(X, \mathbb{C}))$.

Exercises for Section 1.1

Exercise 1.1.1. Let G be a group, endowed with a topology τ. Show that (G, τ) is a topological group if the following conditions are satisfied:

(i) The left multiplication maps $\lambda_g \colon G \to G, x \mapsto gx$ are continuous.

(ii) The inversion map $\eta_G \colon G \to G, x \mapsto x^{-1}$ is continuous.

(iii) The multiplication $m_G \colon G \times G \to G$ is continuous in $(\mathbf{1}, \mathbf{1})$.

Hint: Use (i) and (ii) to derive that all right multiplications and hence all conjugations are continuous.

Exercise 1.1.2. Let G be a group, endowed with a topology τ. Show that (G, τ) is a topological group if the following conditions are satisfied:

(i) The left multiplication maps $\lambda_g \colon G \to G, x \mapsto gx$ are continuous.

(ii) The right multiplication maps $\rho_g \colon G \to G, x \mapsto xg$ are continuous.

(iii) The inversion map $\eta_G \colon G \to G$ is continuous in $\mathbf{1}$.

(iv) The multiplication $m_G \colon G \times G \to G$ is continuous in $(\mathbf{1}, \mathbf{1})$.

Exercise 1.1.3. Let $\alpha : G \to H$ be a homomorphism of topological groups.

(1) α is continuous if and only if α is continuous in $\mathbf{1}$.

(2) α is open if and only if the image $\alpha(U)$ of each identity neighborhood U in G is an identity neighborhood in H.

Exercise 1.1.4. Let $\alpha : G \to H$ be a bijective homomorphism of topological groups for which there exists an identity neighborhood U in G which is mapped homeomorphically on an identity neighborhood $V := \alpha(U)$ in H. Then α is a homeomorphism.

Exercise 1.1.5. Show that if $(G_i)_{i \in I}$ is a family of topological groups, then the product group $G := \prod_{i \in I} G_i$ is a topological group with respect to the product topology.

Exercise 1.1.6. Let G and N be topological groups and suppose that the homomorphism $\alpha \colon G \to \mathrm{Aut}(N)$ defines a continuous map

$$G \times N \to N, \quad (g, n) \mapsto \alpha(g)(n).$$

Then $N \times G$ is a group with respect to the multiplication

$$(n, g)(n', g') := (n\alpha(g)(n'), gg'),$$

called the semidirect product of N and G with respect to α. It is denoted $N \rtimes_\alpha G$. Show that it is a topological group with respect to the product topology.

A typical example is the group

$$\mathrm{Mot}(E) := E \rtimes_\alpha \mathrm{O}(E)$$

of affine isometries of a euclidean space E; also called the *motion group*. In this case $\alpha(g)(v) = gv$ and $\mathrm{Mot}(E)$ acts on E by $(b, g).v := b + gv$ (hence the name).

Exercise 1.1.7. Let X be a compact space and $\mathcal{A}$ be a Banach algebra. Show that:

(a) The space $C(X, \mathcal{A})$ of $\mathcal{A}$-valued continuous functions on X is a complex associative algebra with respect to pointwise multiplication $(fg)(x) := f(x)g(x)$.

(b) $\|f\| := \sup_{x \in X} \|f(x)\|$ is a submultiplicative norm on $C(X, \mathcal{A})$ for which $C(X, \mathcal{A})$ is complete, hence a Banach algebra. Hint: Continuous functions on compact spaces are bounded and uniform limits of sequences of continuous functions are continuous.

(c) $C(X, \mathcal{A})^\times = C(X, \mathcal{A}^\times)$.

Exercise 1.1.8. Let $\mathcal{A}$ be a complex Banach algebra over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. If $\mathcal{A}$ has no unit, we cannot directly associate a "unit group" to $\mathcal{A}$. However, there is a different way to do that by considering on $\mathcal{A}$ the multiplication

$$x * y := x + y + xy.$$

Show that:

(a) The space $\mathcal{A}_+ := \mathcal{A} \times \mathbb{K}$ is a unital Banach algebra with respect to the multiplication

$$(a,t)(a',t') := (aa' + ta' + t'a, tt').$$

(b) The map $\eta\colon \mathcal{A} \to \mathcal{A}_+, x \mapsto (x,1)$ is injective and satisfies $\eta(x * y) = \eta(x)\eta(y)$. Conclude in particular that $(\mathcal{A}, *, 0)$ is a monoid, i.e., a semigroup with neutral element 0.

(c) An element $a \in \mathcal{A}$ is said to be *quasi-invertible* if it is an invertible element in the monoid $(\mathcal{A}, *, 0)$. Show that the set $\mathcal{A}^\times$ of quasi-invertible elements of $\mathcal{A}$ is an open subset and that $(\mathcal{A}^\times, *, 0)$ is a topological group.

Exercise 1.1.9. If $q\colon G \to H$ is a surjective open morphism of topological groups, then the induced map $G/\ker q \to H$ is an isomorphism of topological groups, where $G/\ker q$ is endowed with the quotient topology.

1.2 Subgroups

Throughout this section, G denotes a group and $\mathbf{1}$ its identity element. We call a subset $W \subseteq G$ *symmetric* if $W = W^{-1} := \{w^{-1}\colon w \in W\}$. For two subsets $A, B \subseteq G$ we write $AB := \{ab\colon a \in A, b \in B\}$ for the *product of the sets* A *and* B.

Lemma 1.2.1. *Let G be a topological group. Then the following assertions hold:*

(i) *Let K be a compact and V an open subset of G with $K \subseteq V$. Then there exists an open $U \in \mathfrak{U}_G(\mathbf{1})$ with $KU \subseteq V$.*

(ii) *For each $U \in \mathfrak{U}_G(\mathbf{1})$ and $n \in \mathbb{N}$ there exists a symmetric $W \in \mathfrak{U}_G(\mathbf{1})$ with $W^n \subseteq U$.*

(iii) *If $U \subseteq G$ is open and $M \subseteq G$, then MU and UM are open subsets of G.*

(iv) *If $A, B \subseteq G$ are arcwise connected subsets, then AB is arcwise connected.*

(v) *For a subset $S \subseteq G$ we have*

$$S^0 = \{s \in S\colon (\exists U \in \mathfrak{U}_G(\mathbf{1}))\, sU \subseteq S\} \quad \textit{and} \quad \overline{S} = \bigcap\{SU\colon U \in \mathfrak{U}_G(\mathbf{1})\}.$$

Proof. (i) Let $m_G\colon G \times G \to G$ denote the group multiplication. Then $m_G^{-1}(V) \subseteq G \times G$ is an open subset containing $K \times \{\mathbf{1}\}$. Hence there exists an open $U \in \mathfrak{U}_G(\mathbf{1})$ with $K \times U \subseteq m_G^{-1}(V)$ (Lemma A.4.6), and this means $KU \subseteq V$.

(ii) Since inversion is a homeomorphism, for each $U \in \mathfrak{U}_G(\mathbf{1})$ the set U^{-1} also is a neighborhood of $\mathbf{1}$ and therefore $V := U \cap U^{-1} \in \mathfrak{U}_G(\mathbf{1})$. This means that every $\mathbf{1}$-neighborhood contains a symmetric one.

An easy induction implies that the n-fold multiplication map

$$G^n \to G, \quad (g_1, \ldots, g_n) \to g_1 \cdots g_n$$

is continuous. Hence there exists for each $\mathbf{1}$-neighborhood U a $\mathbf{1}$-neighborhood V with $V^n \subseteq U$. Now $W := V \cap V^{-1}$ is a symmetric $\mathbf{1}$-neighborhood with $W^n \subseteq U$.

(iii) In view of Lemma 1.1.6, the sets $mU \subseteq G$ are open, and MU is the union of these sets, hence open. Likewise UM is open.

(iv) Since A and B are arcwise connected, the same holds for their topological product $A \times B$. Now the continuity of the multiplication map $m_G \colon G \times G \to G$ implies that $AB = m_G(A \times B)$ is arcwise connected (Lemma A.5.3).

(v) In view of Lemma 1.1.6, $s \in S^0$ is equivalent to $s^{-1}S \in \mathfrak{U}_G(\mathbf{1})$. This immediately implies the description of S^0.

For the description of the closure $\overline{S}$, we note that for $x \in G$ the condition $x \in \overline{S}$ is equivalent to $V \cap S \neq \emptyset$ for each $V \in \mathfrak{U}_G(x) = x\mathfrak{U}_G(\mathbf{1})$. This means that $xU \cap S \neq \emptyset$ for each $\mathbf{1}$-neighborhood U, which in turn is $x \in SU^{-1}$. Since η_G is a homeomorphism, for each $\mathbf{1}$-neighborhood U, the set U^{-1} also is a $\mathbf{1}$-neighborhood. Therefore

$$\overline{S} = \bigcap\{SU^{-1} \colon U \in \mathfrak{U}_G(\mathbf{1})\} = \bigcap\{SW \colon W \in \mathfrak{U}_G(\mathbf{1})\}. \qquad \square$$

Closed subgroups

A subset S of a topological space X is called *locally closed* if for each $s \in S$ there exists a neighborhood $U \in \mathfrak{U}_X(s)$ for which $U \cap S$ is a closed subset of U.

Lemma 1.2.2. *Let H be a subgroup of the topological group G.*

(i) *$\overline{H}$ is a subgroup of G.*

(ii) *If H is locally closed, then it is closed.*

(iii) *If H is open, then it is closed.*

(iv) *For each symmetric $\mathbf{1}$-neighborhood $U \in \mathfrak{U}_G(\mathbf{1})$ the set $\langle U\rangle := \bigcup_{n\in\mathbb{N}} U^n$ is an open subgroup of G.*

Proof. (i) Since the map $\varphi \colon G \times G \to G, (x, y) \mapsto xy^{-1}$ is continuous, we obtain

$$\overline{H} \cdot \overline{H}^{-1} = \varphi(\overline{H} \times \overline{H}) = \varphi(\overline{H \times H}) \subseteq \overline{\varphi(H \times H)} \subseteq \overline{H}.$$

This implies that $\overline{H}^{-1} \subseteq \overline{H}$ and hence that $\overline{H}$ is a subgroup of G.

(ii) Let $U \in \mathfrak{U}_G(\mathbf{1})$ be open such that $U \cap H$ is a closed subset of U. Further let $x \in \overline{H}$. Then $x \in HU^{-1}$, so that there exists a $u \in U$ with $y := xu \in H$. Then $u = x^{-1}y \in U$ and $u \in \overline{H} \cdot H = \overline{H}$, so that

$$u \in \overline{H} \cap U = \overline{U \cap H} \cap U = U \cap H$$

implies $x = yu^{-1} \in H$.

(iii) The complement of H is the union of the cosets gH, $g \notin H$. Since all the cosets gH are open (Lemma 1.1.6(ii)), the subgroup H is closed.

(iv) From the symmetry of U we derive $(U^n)^{-1} = (U^{-1})^n = U^n$, showing that $\langle U\rangle$ is invariant under inversion. Therefore $U^nU^m \subseteq U^{n+m}$ implies that $\langle U\rangle$ is a subgroup of G. Since $\langle U\rangle$ contains an open $\mathbf{1}$-neighborhood V, we have $\langle U\rangle = V\langle U\rangle$, and this set is open by Lemma 1.2.1(iii). $\square$

Remark 1.2.3. A subgroup Γ of a topological group G said to be *discrete* if it is discrete as a topological subspace. It is easy to see that this is equivalent to the existence of a $\mathbf{1}$-neighborhood $U \subseteq G$ with $U \cap \Gamma = \{\mathbf{1}\}$. In particular, Γ is locally closed, hence closed by the preceding lemma.

Example 1.2.4. We consider the topological group $G = (\mathbb{R}, +)$. Suppose that $\{0\} \neq \Gamma \subseteq \mathbb{R}$ is a subgroup. Then two cases occur:

Case 1: $\inf(\mathbb{R}_+^\times \cap \Gamma) = 0$, i.e., there exists a sequence $0 < x_n \in \Gamma$ with $x_n \to 0$. Then $\mathbb{Z}x_n \subseteq \Gamma$ holds for each n. For each open interval $]a, b[\subseteq \mathbb{R}$ and $x_n < b - a$ we then obtain

$$\emptyset \neq \mathbb{Z}x_n \cap]a, b[\subseteq \Gamma \cap]a, b[,$$

so that Γ is dense, i.e., $\overline{\Gamma} = \mathbb{R}$.

Case 2: $d := \inf(\mathbb{R}_+^\times \cap \Gamma) > 0$. Then $]-d, d[\cap \Gamma = \{0\}$ implies that Γ is discrete and therefore closed. If $d \notin \Gamma$, then there exists a $d' \in]d, 2d[\cap \Gamma$ and likewise a $d'' \in]d, d'[\cap \Gamma$. Then $0 < d' - d'' < d$ contradicts the definition of d. This implies that $d \in \Gamma$, and hence that $\mathbb{Z}d \subseteq \Gamma$. To see that we actually have equality, let $\gamma \in \Gamma$ and $k := \max\{n \in \mathbb{Z} \colon nd \leq \gamma\}$. Then $\gamma - nd \in [0, d[\cap \Gamma = \{0\}$ implies $\gamma = nd$. We conclude that $\Gamma = \mathbb{Z}d$ is a cyclic group.

In particular, we have shown that all non-trivial closed subgroups of $\mathbb{R}$ are cyclic and isomorphic to $\mathbb{Z}$.

Lemma 1.2.5. *Let $\theta \in \mathbb{R}$. Then $\mathbb{Z} + \mathbb{Z}\theta$ is dense in $\mathbb{R}$ if and only if θ is irrational.*

Proof. Suppose first that $\mathbb{Z}+\mathbb{Z}\theta$ is not dense in $\mathbb{R}$. Then it is discrete by Example 1.2.4, hence of the form $\mathbb{Z}x_o$ for some $x_o > 0$. Then there exist $k, m \in \mathbb{Z}$ with

$$1 = kx_o \quad \text{and} \quad \theta = mx_o.$$

We then obtain $\theta = \frac{m}{k} \in \mathbb{Q}$. If, conversely, $\theta = \frac{m}{k} \in \mathbb{Q}$, then $\mathbb{Z} + \mathbb{Z}\theta \subseteq \frac{1}{k}\mathbb{Z}$ is not dense in $\mathbb{R}$. □

The dense wind

In this short subsection we discuss an important example of a subgroup of the 2-torus $\mathbb{T}^2$ which is not closed. It is the simplest example of a non-closed, arcwise connected subgroup.

Let

$$A = \left\{ \begin{pmatrix} e^{it\sqrt{2}} & 0 \\ 0 & e^{it} \end{pmatrix} : t \in \mathbb{R} \right\} \subseteq \mathbb{T}^2 := \left\{ \begin{pmatrix} e^{ir} & 0 \\ 0 & e^{is} \end{pmatrix} : r, s \in \mathbb{R} \right\},$$

where $\mathbb{T}^2$ is the *two-dimensional torus.* We endow $\mathbb{T}^2$ with the subspace topology inherited from $M_2(\mathbb{C})$.

Lemma 1.2.6. *A is a dense subgroup of the 2-torus $\mathbb{T}^2$.*

Proof. We consider the map

$$\Phi\colon \mathbb{R}^2 \to \mathbb{T}^2, \quad (r,s) \mapsto \begin{pmatrix} e^{2\pi i r} & 0 \\ 0 & e^{2\pi i s} \end{pmatrix}$$

which is a surjective continuous group homomorphism with kernel $\mathbb{Z}^2$. For $L := \mathbb{R}(\sqrt{2}, 1)$ and $V = \mathbb{R}(1,0)$ we have $\mathbb{R}^2 \cong V \oplus L$. In view of

$$A = \Phi(L) = \Phi(L + \mathbb{Z}^2),$$

it suffices to show that $L + \mathbb{Z}^2$ is dense in $\mathbb{R}^2$. From the direct decomposition $\mathbb{R}^2 \cong V \oplus L$ and $L \subseteq L + \mathbb{Z}^2$ we derive

$$L + \mathbb{Z}^2 = L + ((L + \mathbb{Z}^2) \cap V),$$

and if $p\colon \mathbb{R}^2 \to V$ denote the projection map with kernel L, then

$$(L + \mathbb{Z}^2) \cap V = p(L + \mathbb{Z}^2) = p(\mathbb{Z}^2).$$

It therefore suffices to show that $p(\mathbb{Z}^2)$ is dense in V. From $p(1,0) = (1,0)$ and $p(0,1) = p((0,1) - (\sqrt{2},1)) = -(\sqrt{2},0)$ we obtain $p(\mathbb{Z}^2) = \mathbb{Z} + \sqrt{2}\mathbb{Z}$, so that the density of $p(\mathbb{Z}^2)$ is a consequence of Lemma 1.2.5. $\square$

Arc components

Let G be a topological group. We write G_a for the arc component of the identity element $\mathbf{1}$.

Lemma 1.2.7. *G_a is a normal subgroup of G.*

Proof. With Lemma 1.2.1(iv) we see that the product set $G_aG_a \subseteq G$ is arcwise connected. Since it contains $\mathbf{1} = \mathbf{11}$, it follows that $G_aG_a \subseteq G_a$. Moreover, the continuous inversion map $\eta\colon G \to G, g \mapsto g^{-1}$ maps G_a into an arcwise connected set G_a^{-1} which also contains $\mathbf{1} = \mathbf{1}^{-1}$ (Lemma A.5.3). Therefore $G_a^{-1} \subseteq G_a$. Hence G_a is a subgroup of G.

For each $g \in G$ the conjugation automorphism $c_g\colon x \mapsto gxg^{-1}$ is continuous and fixes $\mathbf{1}$. So it maps G_a onto an arcwise connected set containing $\mathbf{1}$. We conclude that $c_g(G_a) \subseteq G_a$. This implies that G_a is a normal subgroup of G. $\square$

Definition 1.2.8. Since G_a is a normal subgroup of G, the set G/G_a of cosets of G_a has a natural group structure. The group

$$\pi_0(G) := G/G_a$$

is called the *component group of G* or the 0-*th homotopy group of G*.

Connected components

Let G be a topological group. We write G_o for the connected component of the identity element $\mathbf{1}$.

Lemma 1.2.9. (i) *G_o is a closed normal subgroup of G.*

(ii) $G_a \subseteq G_o$.

Proof. (i) Similar arguments as in the proof of Lemma 1.2.7, using that continuous maps map connected sets into connected sets (Lemma A.5.3), imply that G_o is a normal subgroup of G. The closedness of G_o follows from the closedness of all connected components of a topological space (Definition A.5.5).

(ii) Finally the connectedness of G_a (Lemma A.5.3(b)) implies that $G_a \subseteq G_o$. □

In all the groups we shall be dealing with, the arc-component G_a will coincide with G_o because G_a is open.

Lemma 1.2.10. *If G_a is open, then $G_a = G_0$.*

Proof. We have already seen that $G_a \subseteq G_o$. Since all cosets gG_a, $g \in G_0$, are open subsets of G_o, the connectedness of G_o implies that this partition is trivial. This means that $G_a = G_o$. □

Exercises for Section 1.2

Exercise 1.2.1. (a) Let $A = A^{-1}$ be a symmetric (arcwise) connected subset of the topological group G containing $\mathbf{1}$. Then

$$H := \langle A \rangle = \bigcup_{n \in \mathbb{N}} A^n$$

is an (arcwise) connected subgroup of G.

(b) Show that the assumption $\mathbf{1} \in A$ in (a) necessary for the arcwise connectedness of H. Consider the subset $A := \{g \in \mathrm{O}_2(\mathbb{R})\colon \det(g) = -1\} \subseteq G = \mathrm{O}_2(\mathbb{R})$.

Exercise 1.2.2. If G is a connected topological group and $H \subseteq G$ an open subgroup, then $G = H$.

Exercise 1.2.3. Let $\alpha : G \to H$ be a morphism of locally compact groups and assume that there exists a relatively compact identity neighborhood U in G for which $\alpha(U)$ is an identity neighborhood in H. Then $\alpha(G)$ is an open subgroup of H and the corestriction map $\alpha : G \to \alpha(G)$ is a homeomorphism.

Exercise 1.2.4. If G is a locally compact group, G_o its identity component and the group G/G_o is countable, then G is the union of countably many compact subsets (this property is called *σ-compactness*).

Exercise 1.2.5. A subset S of the topological space X is locally closed if and only if there exists a closed subset C and an open subset O with $S = C \cap O$.

Exercise 1.2.6. A morphism $\varphi\colon G \to H$ of topological groups is called an *epimorphism* if for any pair of morphisms $f_1, f_2\colon H \to H'$ of topological groups the condition $f_1 \circ \varphi = f_2 \circ \varphi$ implies $f_1 = f_2$.

(1) Show that if φ has dense range in H, then φ is an epimorphism.

(2) Find an example of an epimorphism which is not surjective.

Exercise 1.2.7. Let $D \subseteq \mathbb{R}^n$ be a discrete subgroup. Then there exist linearly independent elements $v_1, \ldots, v_k \in \mathbb{R}^n$ with $D = \sum_{i=1}^k \mathbb{Z}v_i$. Hint: Use induction on $\dim \operatorname{span} D$. If $n > 1$, and D spans $\mathbb{R}^n$, then pick linearly independent elements $f_1, \ldots, f_{n-1} \in D$ and apply induction on $F \cap D$ for $F := \operatorname{span}\{f_1, \ldots, f_{n-1}\}$, where F is a hyperplane in $\mathbb{R}^n$. Now choose $f_n \in D$ with $D = \mathbb{Z}f_n + D \cap F$. This can be done by assuming that $F = \mathbb{R}^{n-1}$ and then choosing f_n with minimal positive nth component (Verify the existence!).

Exercise 1.2.8. Let G be a connected topological group and $\Gamma \trianglelefteq G$ a discrete normal subgroup. Then Γ is central.

1.3 Some Concrete Examples

Definition 1.3.1. We introduce the following notation for some important subgroups of $\mathrm{GL}_n(\mathbb{K})$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$:

(1) The *special linear group* : $\mathrm{SL}_n(\mathbb{K}) := \{g \in \mathrm{GL}_n(\mathbb{K})\colon \det g = 1\}$.

(2) The *orthogonal group* : $\mathrm{O}_n(\mathbb{K}) := \{g \in \mathrm{GL}_n(\mathbb{K})\colon g^\top = g^{-1}\}$.

(3) The *special orthogonal group* : $\mathrm{SO}_n(\mathbb{K}) := \mathrm{SL}_n(\mathbb{K}) \cap \mathrm{O}_n(\mathbb{K})$.

(4) The *unitary group* : $\mathrm{U}_n(\mathbb{K}) := \{g \in \mathrm{GL}_n(\mathbb{K})\colon g^* = g^{-1}\}$. Note that $\mathrm{U}_n(\mathbb{R}) = \mathrm{O}_n(\mathbb{R})$, but $\mathrm{O}_n(\mathbb{C}) \neq \mathrm{U}_n(\mathbb{C})$.

(5) The *special unitary group* : $\mathrm{SU}_n(\mathbb{K}) := \mathrm{SL}_n(\mathbb{K}) \cap \mathrm{U}_n(\mathbb{K})$.

To see that these sets are indeed subgroups, one simply has to use that $(ab)^\top = b^\top a^\top$, $\overline{ab} = \overline{a}\overline{b}$ and that

$$\det\colon \mathrm{GL}_n(\mathbb{K}) \to (\mathbb{K}^\times, \cdot)$$

is a group homomorphism.

Lemma 1.3.2. (a) *The groups* $\mathrm{U}_n(\mathbb{C}), \mathrm{SU}_n(\mathbb{C}), \mathrm{O}_n(\mathbb{R})$ *and* $\mathrm{SO}_n(\mathbb{R})$ *are compact.*
(b) *The groups* $\mathrm{SL}_n(\mathbb{K})$ *and* $\mathrm{O}_n(\mathbb{C})$ *are non-compact for* $n \geq 2$.

Proof. (a) Since all these groups are subsets of $M_n(\mathbb{C}) \cong \mathbb{C}^{n^2}$, we have to show that they are closed and bounded.

Bounded: In view of

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

it suffices to see that $\mathrm{U}_n(\mathbb{C})$ is bounded. Let $g_1, \ldots, g_n$ denote the rows of the matrix $g \in M_n(\mathbb{C})$. Then $g^* = g^{-1}$ is equivalent to $gg^* = \mathbf{1}$, which means that $g_1, \ldots, g_n$ form an orthonormal basis for $\mathbb{C}^n$ with respect to the scalar product $\langle z, w\rangle = \sum_{j=1}^n z_j \overline{w_j}$ which induces the norm $\|z\| = \sqrt{\langle z, z\rangle}$. Therefore $g \in \mathrm{U}_n(\mathbb{C})$ implies $\|g_j\| = 1$ for each j, so that $\mathrm{U}_n(\mathbb{C})$ is bounded.

Closed: The functions

$$f, h\colon M_n(\mathbb{K}) \to M_n(\mathbb{K}), \quad f(A) := AA^* - \mathbf{1} \quad \text{and} \quad h(A) := AA^\top - \mathbf{1}$$

are continuous. Therefore the groups

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

are closed. Likewise $\mathrm{SL}_n(\mathbb{K})$ is closed, and therefore the groups $\mathrm{SU}_n(\mathbb{C})$ and $\mathrm{SO}_n(\mathbb{R})$ are also closed because they are intersections of closed subsets.

(b) Since $\mathrm{SL}_2(\mathbb{R}) \subseteq \mathrm{SL}_2(\mathbb{K}) \subseteq \mathrm{SL}_n(\mathbb{K})$ and $\mathrm{O}_2(\mathbb{C}) \subseteq \mathrm{O}_n(\mathbb{C})$, we may assume that $n = 2$ and show that $\mathrm{SL}_2(\mathbb{R})$ and $\mathrm{O}_2(\mathbb{C})$ are unbounded.

For $\mathrm{SL}_2(\mathbb{R})$ this follows from

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}), \quad \text{for} \quad x \in \mathbb{R},$$

and for $\mathrm{O}_2(\mathbb{C})$ this follows from

$$\begin{pmatrix} \cosh(t) & -i\sinh(t) \\ i\sinh(t) & \cosh(t) \end{pmatrix} \in \mathrm{O}_2(\mathbb{C}) \quad \text{for} \quad t \in \mathbb{R}. \qquad \square$$

.

Proposition 1.3.3. (a) *The group* $\mathrm{U}_n(\mathbb{C})$ *is arcwise connected.*

(b) *The group* $\mathrm{O}_n(\mathbb{R})$ *has the two arc components*

$$\mathrm{SO}_n(\mathbb{R}) \quad \text{and} \quad \mathrm{O}_n(\mathbb{R})_- := \{g \in \mathrm{O}_n(\mathbb{R})\colon \det g = -1\}.$$

Proof. (a) First we consider $\mathrm{U}_n(\mathbb{C})$. To see that this group is arcwise connected, let $g \in \mathrm{U}_n(\mathbb{C})$. Then there exists an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of g. Let $\lambda_1, \ldots, \lambda_n$ denote the corresponding eigenvalues. We write $u := (v_1, \ldots, v_n) \in \mathrm{U}_n(\mathbb{C})$ for the matrix whose columns are $v_1, \ldots, v_n$. Then $u^{-1}gu = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and the unitarity of g implies that $|\lambda_j| = 1$, so that we find $\theta_j \in \mathbb{R}$ with $\lambda_j = e^{\theta_j i}$. Now we define a continuous curve

$$\gamma\colon [0,1] \to \mathrm{U}_n(\mathbb{C}), \quad \gamma(t) := u\,\mathrm{diag}(e^{it\theta_1 i}, \ldots, e^{it\theta_n})u^{-1}.$$

We then have $\gamma(0) = \mathbf{1}$, $\gamma(1) = g$, and each $\gamma(t)$ is unitary.

(b) For $g \in \mathrm{O}_n(\mathbb{R})$ we have $gg^\top = \mathbf{1}$ and therefore $1 = \det(gg^\top) = (\det g)^2$. This shows that

$$\mathrm{O}_n(\mathbb{R}) = \mathrm{SO}_n(\mathbb{R}) \dot\cup \mathrm{O}_n(\mathbb{R})_-$$

and both sets are closed in $\mathrm{O}_n(\mathbb{R})$ because det is continuous. Therefore $\mathrm{O}_n(\mathbb{R})$ is not connected and hence not arcwise connected. If we show that $\mathrm{SO}_n(\mathbb{R})$ is arcwise connected and $x, y \in \mathrm{O}_n(\mathbb{R})_-$, then $\mathbf{1}, x^{-1}y \in \mathrm{SO}_n(\mathbb{R})$ can be connected by an arc $\gamma\colon [0,1] \to \mathrm{SO}_n(\mathbb{R})$, and then $t \mapsto x\gamma(t)$ defines an arc $[0,1] \to \mathrm{O}_n(\mathbb{R})_-$ connecting x to y. So it remains to show that $\mathrm{SO}_n(\mathbb{R})$ is arcwise connected.

Let $g \in \mathrm{SO}_n(\mathbb{R})$. From Linear Algebra we know that there exists an orthogonal matrix u such that ugu^{-1} has the form

$$\begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 \\ \sin\alpha_1 & \cos\alpha_1 \\ & & \ddots \\ & & & \cos\alpha_m & -\sin\alpha_m \\ & & & \sin\alpha_m & \cos\alpha_m \\ & & & & & -1 \\ & & & & & & \ddots \\ & & & & & & & -1 \\ & & & & & & & & 1 \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix}.$$

The determinant of each two-dimensional block is 1, so that the determinant is the product of all -1-eigenvalues. Hence their number is even, and we can write each consecutive pair as a block

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\pi & -\sin\pi \\ \sin\pi & \cos\pi \end{pmatrix}.$$

This shows that

$$ugu^{-1} = \begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 \\ \sin\alpha_1 & \cos\alpha_1 \\ & & \ddots \\ & & & \cos\alpha_m & -\sin\alpha_m \\ & & & \sin\alpha_m & \cos\alpha_m \\ & & & & & 1 \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix}.$$

Now we obtain an arc $\gamma\colon [0,1] \to \mathrm{SO}_n(\mathbb{R})$ with $\gamma(0) = \mathbf{1}$ and $\gamma(1) = g$ by

$$\gamma(t) := u^{-1} \begin{pmatrix} \cos t\alpha_1 & -\sin t\alpha_1 \\ \sin t\alpha_1 & \cos t\alpha_1 \\ & & \ddots \\ & & & \cos t\alpha_m & -\sin t\alpha_m \\ & & & \sin t\alpha_m & \cos t\alpha_m \\ & & & & & 1 \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix} u.$$

□

1.4 From Local Data to Group Topologies

Lemma 1.2.1(v) implies in particular that a subset $O \subseteq G$ is open if and only if for each $g \in O$ the set $g^{-1}O$ is a neighborhood of the identity element e. Let $\mathfrak{U} := \mathfrak{U}_G(\mathbf{1})$ denote the set of all neighborhoods of $\mathbf{1}$. Then the considerations from above imply that the topology τ on G is defined by

$$O \subseteq G \text{ open} \quad \Longleftrightarrow \quad (\forall g \in O)(\exists U \in \mathfrak{U}(\mathbf{1})) \;\; gU \subseteq O. \tag{1.1}$$

This means that the topology on G is completely determined by the set $\mathfrak{U}$. In this sense we think of a topological group as a structure consisting of a group G and an additional structure encoded in the system $\mathfrak{U}$ of $\mathbf{1}$-neighborhoods. Our next step is to make it more precise which systems of subsets occur as $\mathfrak{U}$ for a group topology on G.

The following lemma describes how to construct a *group topology* on a group G, i.e., a Hausdorff topology for which the group multiplication and the inversion are continuous, from a filter basis of subsets which then becomes a filter basis of identity neighborhoods for the group topology.

Definition 1.4.1. Let X be a set. A set $\mathfrak{F} \subseteq \mathbb{P}(X)$ of subsets of X is called a *filter basis* if the following conditions are satisfied:

(F1) $\mathfrak{F} \neq \emptyset$.

(F2) Each set $F \in \mathfrak{F}$ is nonempty.

(F3) $A, B \in \mathfrak{F} \Rightarrow (\exists C \in \mathfrak{F})\; C \subseteq A \cap B$.

Example 1.4.2. (a) If (X, τ) is a topological space and $x \in X$, then the set $\mathfrak{U}_X(x) = \mathfrak{U}(x)$ of all neighborhoods of x is a filter, called the *neighborhood filter of* x.

Any non-empty system $\mathfrak{B}$ of neighborhoods of x with the property

$$(\forall U \in \mathfrak{U}(x))\,(\exists B \in \mathfrak{B}) \quad B \subseteq U$$

is called a *basis of neighborhoods of* x. This implies that $\mathfrak{B}$ is a filter basis generating the filter $\mathfrak{U}(x)$.

Lemma 1.4.3. *Let G be a topological group and $\mathcal{F}$ be a filter basis of $\mathfrak{U}_G(e)$. Then*

(U0) $\bigcap \mathfrak{F} = \{\mathbf{1}\}$.

(U1) $(\forall U \in \mathfrak{F})(\exists V \in \mathfrak{F})\; VV \subseteq U$.

(U2) $(\forall U \in \mathfrak{F})(\exists V \in \mathfrak{F})\; V^{-1} \subseteq U$.

(U3) $(\forall U \in \mathfrak{F})(\forall g \in G)(\exists V \in \mathfrak{F})\; gVg^{-1} \subseteq U$.

Proof. (U0) follows from the fact that G is separated.

(U1) follows from the continuity of the multiplication map in $(\mathbf{1}, \mathbf{1})$.

(U2) follows from the continuity of the inversion map in $\mathbf{1}$.

(U3) follows from the continuity of the conjugation map c_g in $\mathbf{1}$. □

Lemma 1.4.4. *Let G be a group and $\mathfrak{F}$ a filter basis of subsets of G satisfying*

(U0) $\bigcap \mathcal{F} = \{\mathbf{1}\}$.

(U1) $(\forall U \in \mathfrak{F})(\exists V \in \mathfrak{F})\ VV \subseteq U$.

(U2) $(\forall U \in \mathfrak{F})(\exists V \in \mathfrak{F})\ V^{-1} \subseteq U$.

(U3) $(\forall U \in \mathfrak{F})(\forall g \in G)(\exists V \in \mathfrak{F})\ gVg^{-1} \subseteq U$.

Then there exists a unique group topology τ on G such that $\mathfrak{F}$ is a basis of $\mathbf{1}$-neighborhoods in G. This topology is given by

$$\tau = \{O \subseteq G \colon (\forall g \in O)(\exists V \in \mathfrak{F})\ gV \subseteq O\}.$$

Proof. First we show that τ is a topology. Clearly $\emptyset, G \in \tau$. Let $(U_j)_{j \in J}$ be a family of elements of τ and $U := \bigcup_{j \in J} U_j$. For each $g \in U$, there exists a $j_0 \in J$ with $g \in U_{j_0}$ and a $V \in \mathfrak{F}$ with $gV \subseteq U_{j_0} \subseteq U$. Thus $U \in \tau$ and we see that τ is stable under arbitrary unions.

If $U_1, U_2 \in \tau$ and $g \in U_1 \cap U_2$, there exist $V_1, V_2 \in \mathfrak{F}$ with $gV_i \subseteq U_i$. Since $\mathcal{F}$ is a filter basis, there exists $V_3 \in \mathcal{F}$ with $V_3 \subseteq V_1 \cap V_2$, and then $gV_3 \subseteq U_1 \cap U_2$. We conclude that $U_1 \cap U_2 \in \tau$, and hence that τ is a topology on G.

We claim that the interior U° of a subset $U \subseteq G$ is given by

$$U_1 := \{u \in U \colon (\exists V \in \mathcal{F})\ uV \subseteq U\}.$$

In fact, if there exists a $V \in \mathcal{F}$ with $uV \subseteq U$, then we pick a $W \in \mathcal{F}$ with $WW \subseteq V$ and obtain $uWW \subseteq U$, so that $uW \subseteq U_1$. Hence $U_1 \in \tau$, i.e., U_1 is open, and it clearly is the largest open subset contained in U, i.e., $U_1 = U^\circ$. It follows in particular that U is a neighborhood of g if and only if $g \in U^\circ$, and we see in particular that $\mathcal{F}$ is a neighborhood basis at $\mathbf{1}$. The property $\bigcap \mathcal{F} = \{\mathbf{1}\}$ implies that for $x \neq y$ there exists $U \in \mathcal{F}$ with $y^{-1}x \notin U$. For $V \in \mathcal{F}$ with $VV \subseteq U$ and $W \in \mathcal{F}$ with $W^{-1} \subseteq V$ we then obtain $y^{-1}x \notin VW^{-1}$, i.e., $xW \cap yV = \emptyset$. Thus (G, τ) is a Hausdorff space.

To see that G is a topological group, we have to verify that the map

$$f \colon G \times G \to G, \quad (x, y) \mapsto xy^{-1}$$

is continuous. So let $x, y \in G$, $U \in \mathcal{F}$ and pick $V \in \mathcal{F}$ with $yVy^{-1} \subseteq U$ and $W \in \mathcal{F}$ with $WW^{-1} \subseteq V$ (cf. (U2/3)). Then

$$f(xW, yW) = xWW^{-1}y^{-1} = xy^{-1}y(WW^{-1})y^{-1} \subseteq xy^{-1}yVy^{-1} \subseteq xy^{-1}U$$

implies that f is continuous in (x, y).

The preceding arguments show that τ is a group topology on G for which $\mathcal{F}$ is a basis of $\mathbf{1}$-neighborhoods. That τ is uniquely determined by this property follows from (1.1). ☐

Lemma 1.4.5. *Let G be a group and $U = U^{-1}$ a symmetric subset containing $\mathbf{1}$. We further assume that U carries a Hausdorff topology for which*

(T1) *$D := \{(x,y) \in U \times U : xy \in U\}$ is an open subset of $U \times U$ and the group multiplication $m_U : D \to U, (x,y) \mapsto xy$ is continuous,*

(T2) *the inversion map $\iota_U : U \to U, u \mapsto u^{-1}$ is continuous, and*

(T3) *for each $g \in G$, there exists an open $\mathbf{1}$-neighborhood U_g in U with $c_g(U_g) \subseteq U$, such that the conjugation map $c_g : U_g \to U, x \mapsto gxg^{-1}$ is continuous.*

Then there exists a unique group topology on G for which the inclusion map $U \hookrightarrow G$ is a homeomorphism onto an open subset of G.

If, in addition, U generates G, then (T1/2) *imply* (T3).

Proof. First we consider the filter basis $\mathfrak{F}$ of $\mathbf{1}$-neighborhoods in U. Then (T1) implies (U1), (T2) implies (U2), and (T3) implies (U3). Moreover, the assumption that U is Hausdorff implies that $\bigcap \mathfrak{F} = \{\mathbf{1}\}$. Therefore Lemma 1.4.4 implies that G carries a unique structure of a (Hausdorff) topological group for which $\mathfrak{F}$ is a basis of $\mathbf{1}$-neighborhoods.

We claim that the inclusion map $U \to G$ is an open embedding. So let $x \in U$. Then

$$U_x := U \cap x^{-1}U = \{y \in U : (x,y) \in D\}$$

is open in U and λ_x restricts to a continuous map $U_x \to U$ with image $U_{x^{-1}}$. Its inverse is also continuous. Hence $\lambda_x^U : U_x \to U_{x^{-1}}$ is a homeomorphism. We conclude that the sets of the form xV, where V a neighborhood of $\mathbf{1}$, form a basis of neighborhoods of x in the topological space U. Hence the inclusion map $U \hookrightarrow G$ is an open embedding.

Suppose, in addition, that G is generated by U. For each $g \in U$, there exists an open $\mathbf{1}$-neighborhood U_g with $gU_g \times \{g^{-1}\} \subseteq D$. Then $c_g(U_g) \subseteq U$, and the continuity of m_U implies that $c_g|_{U_g} : U_g \to U$ is continuous.

Hence, for each $g \in U$, the conjugation c_g is continuous in a neighborhood of $\mathbf{1}$. Since the set of all these g is a submonoid of G containing U, it contains U^n for each $n \in \mathbb{N}$, hence all of G because G is generated by $U = U^{-1}$. Therefore (T3) follows from (T1) and (T2). □

Quotient groups

Proposition 1.4.6. *Let G be a topological group and $H \trianglelefteq G$ a closed normal subgroup. Then the quotient topology turns G/H into a topological group. The quotient homomorphism $q : G \to G/H, g \mapsto gH$ is continuous and open.*

Proof. On G/H we consider the filter basis

$$\mathfrak{F} := \{UH : U \in \mathfrak{U}_G(\mathbf{1})\}$$

and verify the conditions (U0)-(U3) from Lemma 1.4.4.

(U0): $\bigcap \mathfrak{F} = \bigcap_{U \in \mathfrak{U}_G(\mathbf{1})} UH = \overline{H} = H$ follows from Lemma 1.2.1.

(U1) follows from $(VH)(VH) = V^2H \subseteq UH$ for $V^2 \subseteq U$.

(U2) follows from $(UH)^{-1} = HU^{-1} = U^{-1}H$.

(U3) follows from $(gH)(UH)(gH)^{-1} = (gUg^{-1})H$ for $g \in G$.

Now Lemma 1.4.4 implies the existence of a unique group topology on G/H for which $\mathfrak{F}$ is a basis of $\mathfrak{U}_{G/H}(\mathbf{1})$.

We consider the quotient homomorphism $q\colon G \to G/H$. Then $q^{-1}(UH)$ contains U for each $U \in \mathfrak{U}_G(\mathbf{1})$, and therefore q is continuous in $\mathbf{1}$, hence continuous (Exercise 1.1.3). Moreover, for each open e-neighborhood $U \subseteq G$ the set

$$q(U) = UH$$

is an identity neighborhood in G/H, showing that q is open (Exercise 1.1.3).

If $O \subseteq G/H$ is an open subset, then $q^{-1}(O)$ is open because q is continuous. If, conversely, $q^{-1}(O)$ is open for a subset $O \subseteq G/H$, then $O = q(q^{-1}(O))$ and the openness of q entail that O is open. Therefore the topology on G/H coincides with the quotient topology with respect to the equivalence relation $x \sim y :\Leftrightarrow xH = yH$. □

Homomorphisms of topological groups

In the following we call a continuous homomorphism $\varphi\colon G \to H$ of topological groups simply a *morphism of topological groups.* These are the mappings between topological groups which are compatible with the topological structure and the group structure. In the same spirit an *isomorphism of topological groups* is a morphism $\varphi\colon G \to H$ for which there exists a morphism $\psi\colon H \to G$ with $\varphi \circ \psi = \mathrm{id}_H$ and $\psi \circ \varphi = \mathrm{id}_G$. This is equivalent to φ being bijective, continuous, and open.

We recall that for each morphism $\varphi\colon G \to H$ of topological groups, the kernel $\ker(\varphi) = \varphi^{-1}(\mathbf{1})$ is a closed normal subgroup, so that the quotient group $G/\ker\varphi$ inherits a natural Hausdorff group topology. Let $\pi\colon G \to G/\ker\varphi, g \mapsto g\cdot\ker\varphi$ denote the quotient map. Then π is open and continuous (Proposition 1.4.6), and φ induces an injective group homomorphism

$$\overline{\varphi}\colon G/\ker\varphi \to H, \quad g\cdot\ker\varphi \mapsto \varphi(g) \quad \text{satisfying} \quad \varphi = \overline{\varphi}\circ\pi.$$

This is called the *canonical factorization of* φ, because it expresses φ as a composition of a surjective open morphism and an injective morphism (Verify the continuity of $\overline{\varphi}$!).

Lemma 1.4.7. *For a surjective morphism $\varphi\colon G \to H$ of topological groups the following are equivalent:*

(i) *φ is open.*

(ii) *The induced injective morphism $\overline{\varphi}\colon G/\ker\varphi \to H, g\ker\varphi \mapsto \varphi(g)$ is an isomorphism of topological groups.*

Proof. Let $\pi\colon G \to G/\ker\varphi$ denote the quotient map and recall from Proposition 1.4.6 that π is an open morphism of topological groups satisfying $\overline{\varphi}\circ\pi = \varphi$. The continuity of the group homomorphism $\overline{\varphi}$ follows from the universal property of the quotient topology on $G/\ker\varphi$ and the continuity of $\varphi = \overline{\varphi}\circ\pi$.

(i) ⇒ (ii): Since $\overline{\varphi}$ is a bijective morphism of topological groups, it suffices to show that $\overline{\varphi}$ is an open map. So let $O \subseteq G/\ker\varphi$ be an open set. Then

$$\overline{\varphi}(O) = \varphi(\pi^{-1}(O))$$

is an open subset of H because φ is open and π is continuous.

(ii) $\Rightarrow$ (i): Since $\varphi = \overline{\varphi} \circ \pi$ is a composition of two open maps, it is an open map. □

Example 1.4.8. (The torus groups) We consider the n-torus

$$\mathbb{T}^n := \{z \in \mathbb{C}^n : (\forall j)|z_j| = 1\},$$

which is a compact abelian group. We have a surjective continuous homomorphism

$$q \colon \mathbb{R}^n \to \mathbb{T}^n, \quad x \mapsto (e^{2\pi i x_1}, \ldots, e^{2\pi i x_n})$$

whose kernel is $\mathbb{Z}^n$. We claim that the induced homomorphism

$$\overline{q} \colon \mathbb{R}^n/\mathbb{Z}^n \to \mathbb{T}^n, \quad \overline{x} = x + \mathbb{Z}^n \mapsto q(x)$$

is a homeomorphism, hence an isomorphism of topological groups. We have already seen above that $\overline{q}$ is continuous. By definition, it is bijective.

Finally we observe that $\mathbb{R}^n/\mathbb{Z}^n = \overline{q}([0,1]^n)$, and since the cube $[0,1]^n$ is compact, the quotient group $\mathbb{R}^n/\mathbb{Z}^n$ is compact. Now our claim follows from the fact that continuous bijections between compact spaces are homeomorphisms (Proposition A.4.4).

Exercises for Section 1.4

Exercise 1.4.1. Let H be a closed subgroup of the topological group G. We endow the coset space G/H with the quotient topology and consider the left multiplication action $\sigma \colon G \times G/H \to G/H, (g, xH) \mapsto gxH$. Show that:

(1) The quotient map $q \colon G \to G/H, g \mapsto gH$ is an open continuous map.

(2) The action $\sigma \colon G \times G/H \to G/H$ is continuous.

Exercise 1.4.2. Let

$$\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} : \|x\|_2 = 1\}$$

denote the n-sphere. Show that

(i) $\sigma(g, x) := gx$ defines a continuous action of $G := \mathrm{O}_{n+1}(\mathbb{R})$ on $\mathbb{S}^n$.

(ii) This action is transitive.

(iii) The stabilizer G_{e_1} of the first basis vector e_1 is isomorphic to $\mathrm{O}_n(\mathbb{R})$.

(iv) The orbit map $\mathrm{O}_{n+1}(\mathbb{R}) \to \mathbb{S}^n, g \mapsto ge_1$ factors through a homeomorphism

$$\mathrm{O}_{n+1}(\mathbb{R})/\mathrm{O}_n(\mathbb{R}) \to \mathbb{S}^n.$$

Chapter 2

The Exponential Function of a Banach Algebra

In this chapter we study one of the central tools in Lie theory: the exponential function of a Banach algebra, the natural generalization of the matrix exponential function. It has various applications in the structure theory of Lie groups. First of all, it is naturally linked to one-parameter subgroups, and it turns out that the local group structure of $\mathcal{A}^\times$ for a unital Banach algebra $\mathcal{A}$ in a neighborhood of the identity is determined by its one-parameter subgroups via the Hausdorff series.

In Section 2.1, we discuss some basic properties of the exponential function of $\mathcal{A}$ and in Section 2.2 we then turn to the logarithm function.

Throughout we shall use the concept of a smooth function on a domain in a Banach space for which we refer to Appendix B.

2.1 Elementary Properties of the Exponential Function

Let $\mathcal{A}$ be a unital Banach algebra. For $x \in \mathcal{A}$ we define

$$e^x := \sum_{k=0}^{\infty} \frac{1}{k!} x^k. \tag{2.1}$$

The absolute convergence of the series on the right follows directly from the estimate

$$\sum_{k=0}^{\infty} \frac{1}{k!} \|x^k\| \leq \sum_{k=0}^{\infty} \frac{1}{k!} \|x\|^k = e^{\|x\|}$$

and the Comparison Test for absolute convergence of a series in a Banach space. We define the *exponential function of* $\mathcal{A}$ by

$$\exp\colon \mathcal{A} \to \mathcal{A}, \quad \exp(x) := e^x.$$

Lemma 2.1.1. *Let $x, y \in \mathcal{A}$.*

(i) *If $xy = yx$, then* $\exp(x + y) = \exp x \exp y$.

(ii) $\exp(\mathcal{A}) \subseteq \mathcal{A}^\times$, $\exp(0) = \mathbf{1}$, *and* $(\exp x)^{-1} = \exp(-x)$.

(iii) *For $g \in \mathcal{A}^\times$ we have $ge^x g^{-1} = e^{gxg^{-1}}$.*

(iv) $\exp$ *is smooth.*

(v) $\mathtt{d}\exp(0) = \mathrm{id}_{\mathcal{A}}$, *and for $xy = yx$ we have*

$$\mathtt{d}\exp(x)y = \exp(x)y = y\exp(x). \tag{2.2}$$

Proof. (i) Using the general form of the Cauchy Product Formula (Exercise 2.1.2), we obtain

$$\begin{aligned}\exp(x+y) &= \sum_{k=0}^{\infty} \frac{(x+y)^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\ell=0}^{k} \binom{k}{\ell} x^\ell y^{k-\ell} \\ &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{k} \frac{x^\ell}{\ell!} \frac{y^{k-\ell}}{(k-\ell)!} = \Big(\sum_{p=0}^{\infty} \frac{x^p}{p!}\Big)\Big(\sum_{\ell=0}^{\infty} \frac{y^\ell}{\ell!}\Big).\end{aligned}$$

(ii) From (i) we derive in particular $\exp x \exp(-x) = \exp 0 = \mathbf{1}$, which implies (ii).

(iii) is a consequence of $gx^n g^{-1} = (gxg^{-1})^n$ and the continuity of the conjugation map $c_g(x) := gxg^{-1}$ on $\mathcal{A}$.

(iv) For this point we shall use the tools from Appendix B. We write the exponential function as $\exp(x) = \sum_{n=0}^{\infty} c_n(x, \ldots, x)$ for

$$c_n(x_1, \ldots, x_n) := \frac{1}{n!} x_1 \cdots x_n.$$

Then each $c_n \colon \mathcal{A}^n \to \mathcal{A}$ is a continuous n-linear function with $\|c_n\| \leq 1/n!$. In particular, $\sum_n \|c_n\| r^n$ converges for every $r > 0$, so that Theorem B.3.7 implies that exp is smooth with

$$\mathtt{d}\exp(0) = \sum_n \mathtt{d}c_n(0).$$

As c_n is n-linear, we have $\mathtt{d}c_n(0) = 0$ for $n > 1$ (cf. Lemma B.2.3), so that the only contribution comes from $c_1(x) = x$ with $\mathtt{d}c_1(0) = c_1 = \mathrm{id}_{\mathcal{A}}$.

(v) We have just seen that $\mathtt{d}\exp(0) = \mathrm{id}_{\mathcal{A}}$. Assume that x and y commute. For $t \in \mathbb{R}$, the relation

$$\exp(x + ty) = \exp(x)\exp(ty)$$

now leads to

$$\mathtt{d}\exp(x)y = \exp(x)\mathtt{d}\exp(0)y = \exp(x)y = y\exp(x). \qquad \square$$

Remark 2.1.2. (a) For $n = 1$, the exponential function

$$\exp\colon \mathbb{R} \cong M_n(\mathbb{R}) \to \mathbb{R}^\times \cong \mathrm{GL}_n(\mathbb{R}), \quad x \mapsto e^x$$

is injective, but this is not the case for $n > 1$. In fact,

$$\exp\begin{pmatrix} 0 & -2\pi \\ 2\pi & 0 \end{pmatrix} = \mathbf{1}$$

follows from

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

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

Product and commutator formula

Definition 2.1.3. A *one-parameter (sub)group* of a group G is a group homomorphism $\gamma\colon (\mathbb{R}, +) \to G$. The following result describes the differentiable one-parameter subgroups of $\mathcal{A}^\times$.

Remark 2.1.4. In the proof of the following theorem, we shall need Banach space valued Riemann integrals. The existence of the Riemann integral

$$\int_a^b \alpha(t)\,dt$$

of a continuous curve $\alpha\colon [a, b] \to E$ with values in a Banach space E is proved with exactly the same arguments as for real-valued integrals. If α is a step function, i.e., constant on the intervals $]x_i, x_{i+1}[$ for some partition

$$a = x_0 < x_1 < \ldots < x_n = b$$

and $\xi \in]x_i, x_{i+1}[$, then the integral is simply given by

$$\int_a^b \alpha(t)\,dt = \sum_{i=0}^{n-1} (t_{i+1} - t_i)\alpha(\xi_i).$$

In this case the relation

$$\Big\| \int_a^b \alpha(t)\,dt \Big\| \leq \sum_{i=0}^{n-1} (t_{i+1} - t_i)\|\alpha(\xi_i)\| = \int_a^b \|\alpha(t)\|\,dt$$

follows from the triangle inequality. For general continuous curves, this relation is obtained by passing to the limit on both sides:

$$\Big\| \int_a^b \alpha(t)\,dt \Big\| \leq \int_a^b \|\alpha(t)\|\,dt. \tag{2.3}$$

Theorem 2.1.5. (Automatic smoothness of one-parameter groups) *For each $x \in \mathcal{A}$, the map*

$$\gamma_x\colon (\mathbb{R},+) \to \mathcal{A}, \quad t \mapsto \exp(tx)$$

is a smooth group homomorphism solving the initial value problem

$$\gamma_x(0) = \mathbf{1} \quad \textit{and} \quad \gamma_x'(t) = \gamma_x(t)x \quad \textit{for } t \in \mathbb{R}.$$

Conversely, every continuous one-parameter group $\gamma\colon \mathbb{R} \to \mathcal{A}^\times$ is of this form.

Proof. In view of Lemma 2.1.1(i) and the differentiability of exp in 0, we have

$$\begin{aligned}\lim_{h\to 0} \frac{1}{h}\big(\gamma_x(t+h) - \gamma_x(t)\big) &= \lim_{h\to 0} \frac{1}{h}\big(\gamma_x(t)\gamma_x(h) - \gamma_x(t)\big) \\ &= \gamma_x(t) \lim_{h\to 0} \frac{1}{h}\big(e^{hx} - \mathbf{1}\big) = \gamma_x(t)x.\end{aligned}$$

Hence γ_x is differentiable with $\gamma_x'(t) = x\gamma_x(t) = \gamma_x(t)x$. From that it immediately follows that γ_x is smooth with $\gamma_x^{(n)}(t) = x^n\gamma_x(t)$ for each $n \in \mathbb{N}$.

We first show that each one-parameter group $\gamma\colon \mathbb{R} \to \mathcal{A}^\times$ which is differentiable in 0 has the required form. For $x := \gamma'(0)$, the calculation

$$\gamma'(t) = \lim_{s\to 0} \frac{\gamma(t+s) - \gamma(t)}{s} = \lim_{s\to 0} \gamma(t)\frac{\gamma(s) - \gamma(0)}{s} = \gamma(t)\gamma'(0) = \gamma(t)x$$

implies that γ is continuously differentiable. Therefore

$$\frac{d}{dt}(e^{-tx}\gamma(t)) = -e^{-tx}x\gamma(t) + e^{-tx}\gamma'(t) = 0$$

implies that $e^{-tx}\gamma(t) = \gamma(0) = \mathbf{1}$ for each $t \in \mathbb{R}$, so that $\gamma(t) = e^{tx}$.

Eventually we consider the general case, where $\gamma\colon \mathbb{R} \to \mathcal{A}^\times$ is only assumed to be continuous. The idea is to construct a differentiable function $\widetilde{\gamma}$ by applying a smoothing procedure to γ and to show that the smoothness of $\widetilde{\gamma}$ implies that of γ. So let $f\colon \mathbb{R} \to \mathbb{R}_+$ be a twice continuously differentiable function with $f(t) = 0$ for $|t| > \varepsilon$ and $\int_{\mathbb{R}} f(t)\,dt = 1$, where ε is chosen such that $\|\gamma(t) - \mathbf{1}\| < \frac{1}{2}$ holds for $|t| \leq \varepsilon$.

We define

$$\widetilde{\gamma}(t) := \int_{\mathbb{R}} f(s)\gamma(t-s)\ ds = \gamma(t)\int_{\mathbb{R}} f(s)\gamma(-s)\ ds = \gamma(t)\int_{-\varepsilon}^{\varepsilon} f(s)\gamma(-s)\,ds.$$

Here we use the existence of Riemann integrals of continuous curves with values in Banach spaces (Remark 2.1.4). Change of Variables leads to

$$\widetilde{\gamma}(t) = \int_{\mathbb{R}} f(t-s)\gamma(s)\,ds,$$

which is differentiable because

$$\frac{\widetilde{\gamma}(t+h) - \widetilde{\gamma}(t)}{h} = \int_{\mathbb{R}} \frac{f(t+h-s) - f(t-s)}{h}\gamma(s)\ ds$$

and the functions $f_h(t) := \frac{f(t+h)-f(t)}{h}$ converge uniformly for $h \to 0$ to f' (this is a consequence of the Mean Value Theorem). We further have

$$\Big\|\int_{-\varepsilon}^{\varepsilon} f(s)\gamma(-s)\,ds - \mathbf{1}\Big\| = \Big\|\int_{-\varepsilon}^{\varepsilon} f(s)\big(\gamma(-s)-\mathbf{1}\big)\,ds\Big\|$$
$$\leq \int_{-\varepsilon}^{\varepsilon} f(s)\|\gamma(-s)-\mathbf{1}\|\,ds \leq \tfrac{1}{2}\int_{-\varepsilon}^{\varepsilon} f(s)\;ds = \tfrac{1}{2},$$

because of the inequality $\|\int h(s)\;ds\| \leq \int \|h(s)\|\,ds$ (cf. (2.3) in Remark 2.1.4).

Let $g := \int_{-\varepsilon}^{\varepsilon} f(s)\gamma(-s)\,ds$. In view of $\|g-\mathbf{1}\| \leq \frac{1}{2}$ we have $g \in \mathcal{A}^\times$ (see the proof of Proposition 1.1.9) and therefore $\gamma(t) = \widetilde{\gamma}(t)g^{-1}$. Now the differentiability of $\widetilde{\gamma}$ implies that γ is differentiable, and one can argue as above. □

The local inverse

Proposition 2.1.6. *There exists an open* 0*-neighborhood* U *in* $\mathcal{A}$*, for which he map*

$$\exp|_U \colon U \to \mathcal{A}^\times$$

is a diffeomorphism onto an open neighborhood of $\mathbf{1}$ *in* $\mathcal{A}^\times$*.*

Proof. We have already seen that exp is a smooth map, and that $\mathtt{d}\exp(\mathbf{0}) = \mathrm{id}_{\mathcal{A}}$ (Lemma 2.1.1). Therefore the assertion follows from the Inverse Function Theorem. □

If U is as in Proposition 2.1.6 and $V = \exp(U)$, we define

$$\log_V \colon = (\exp|_U)^{-1} \colon V \to U \subseteq \mathcal{A}.$$

We shall see below why this function deserves to be called a *logarithm function.*

Theorem 2.1.7. (No Small Subgroup Theorem) *There exists an open neighborhood* V *of* $\mathbf{1}$ *in* $\mathcal{A}^\times$ *such that* $\{\mathbf{1}\}$ *is the only subgroup of* $\mathcal{A}^\times$ *contained in* V*.*

Proof. Let U be as in Proposition 2.1.6 and assume further that U is convex and bounded. We set $U_1 := \frac{1}{2}U$ and observe that $V := \exp U_1$ is an open $\mathbf{1}$-neighborhood in $\mathcal{A}$. Let $G \subseteq V$ be a subgroup of $\mathcal{A}^\times$ and $g \in G$. Then we write $g = \exp x$ with $x \in U_1$ and assume that $x \neq 0$. Let $k \in \mathbb{N}$ be maximal with $kx \in U_1$ (the existence of k follows from the boundedness of U). Then

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

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 $g = \mathbf{1}$. □

Exercises for Section 2.1

Exercise 2.1.1. Let Y be a Banach space and $a_{n,m}$, $n,m \in \mathbb{N}$, elements in Y with

$$\sum_{n,m} \|a_{n,m}\| := \sup_{N\in\mathbb{N}} \sum_{n,m\leq N} \|a_{n,m}\| < \infty.$$

(a) Show that

$$A := \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{n,m}$$

and that both iterated sums exist.

(b) Show that for each sequence $(S_n)_{n\in\mathbb{N}}$ of finite subsets $S_n \subseteq \mathbb{N}\times\mathbb{N}$, $n\in\mathbb{N}$, with $S_n \subseteq S_{n+1}$ and $\bigcup_n S_n = \mathbb{N}\times\mathbb{N}$ we have

$$A = \lim_{n\to\infty} \sum_{(j,k)\in S_n} a_{j,k}.$$

Exercise 2.1.2. (Cauchy Product Formula) Let X, Y, Z be Banach spaces and $\beta\colon X\times Y \to Z$ a continuous bilinear map. Suppose that if $x := \sum_{n=0}^{\infty} x_n$ is absolutely convergent in X and if $y := \sum_{n=0}^{\infty} y_n$ is absolutely convergent in Y, then

$$\beta(x,y) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} \beta(x_k, y_{n-k}).$$

Hint: Use Exercise 2.1.1(b).

Exercise 2.1.3. The function

$$\Phi : \mathbb{R}\to\mathbb{R}, \quad t\mapsto \begin{cases} e^{-\frac{1}{t}}, & \text{for } t>0 \\ 0, & \text{for } t\leq 0 \end{cases}$$

is smooth. Hint: The higher derivatives of $e^{-\frac{1}{t}}$ are of the form $P(t^{-1})e^{-\frac{1}{t}}$, where P is a polynomial.

(b) For $\lambda > 0$ the function $\Psi(t) := \Phi(t)\Phi(\lambda - t)$ is a non-negative smooth function with $\mathrm{supp}(\Psi) = [0,\lambda]$.

Exercise 2.1.4. (A smoothing procedure) Let $f \in C^1_c(\mathbb{R})$ be a C^1-function with compact support and $\gamma \in C(\mathbb{R}, E)$, where E is a Banach space. Then the convolution

$$h := f * \gamma\colon \mathbb{R}\to E, \quad t\mapsto \int_{\mathbb{R}} f(s)\gamma(t-s)\,ds = \int_{\mathbb{R}} f(t-s)\gamma(s)\,ds$$

of f and γ is continuously differentiable with $h' = f' * \gamma$. Hint:

$$\int_{\mathbb{R}} f(t-s)\gamma(s)\,ds = \int_{t-\mathrm{supp}(f)} f(t-s)\gamma(s)\,ds.$$

Exercise 2.1.5. Show that for $\mathcal{A} := C(\mathbb{S}^1, \mathbb{C})$ the exponential function

$$\exp : \mathcal{A} \to \mathcal{A}^\times = C(\mathbb{S}^1, \mathbb{C}^\times), \quad a \mapsto e^a$$

is not surjective. It requires some covering theory to determine which elements $f \in C(\mathbb{S}^1, \mathbb{C}^\times)$ lie in its image. Hint: Use the winding number with respect to 0.

Exercise 2.1.6. Show that for the Banach algebra $\mathcal{A} := L^\infty([0,1], \mathbb{C})$, the exponential function

$$\exp : \mathcal{A} \to \mathcal{A}^\times, \quad a \mapsto e^a$$

is surjective.

Exercise 2.1.7. (a) Calculate e^{tN} for $t \in \mathbb{K}$ and the matrix

$$N = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ \cdot & 0 & 1 & 0 & \cdot \\ \cdot & & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & 1 \\ 0 & & \ldots & & 0 \end{pmatrix} \in M_n(\mathbb{K}).$$

(b) If A is a block diagonal matrix $\operatorname{diag}(A_1, \ldots, A_k)$, then e^A is the block diagonal matrix $\operatorname{diag}(e^{A_1}, \ldots, e^{A_k})$.

(c) Calculate e^{tA} for a matrix $A \in M_n(\mathbb{C})$ given in Jordan normal form. Hint: Use (a) and (b).

Exercise 2.1.8. For $A \in M_n(\mathbb{C})$ we have $e^A = \mathbf{1}$ if and only if A is diagonalizable with all eigenvalues contained in $2\pi i\mathbb{Z}$.

Exercise 2.1.9. Let $V \subseteq M_n(\mathbb{C})$ be a commutative subspace, i.e., an abelian Lie subalgebra. Then $A := e^V$ is an abelian subgroup of $\mathrm{GL}_n(\mathbb{C})$ and

$$\exp : (V, +) \to (A, \cdot)$$

is a group homomorphism whose kernel consists of diagonalizable elements whose eigenvalues are contained in $2\pi i\mathbb{Z}$.

Exercise 2.1.10. Let $D \in M_n(\mathbb{K})$ be a diagonal matrix. Calculate its operator norm with respect to the euclidean norm on $\mathbb{K}^n$.

Exercise 2.1.11. Let $A \in M_n(\mathbb{C})$. Show that the set $e^{\mathbb{R}A} = \{e^{tA} : t \in \mathbb{R}\}$ is bounded in $M_n(\mathbb{C})$ if and only if A is diagonalizable with purely imaginary eigenvalues.

Exercise 2.1.12. Let $U \in M_n(\mathbb{C})$. Then the set $\{U^n : n \in \mathbb{Z}\}$ is bounded if and only if U is diagonalizable and $\operatorname{Spec}(U) \subseteq \{z \in \mathbb{C} : |z| = \mathbf{1}\}$.

2.2 The Logarithm Function

To deal with the logarithm function of a Banach algebra, we need some tools to verify identities such as $\exp(\log(x)) = x$. The following proposition provides a natural tool. It shows in particular that inserting elements of a Banach algebra in power series is compatible with composition.

In the following we write $\mathbb{K}[[z]]$ for the space of all formal power series

$$f(z) := \sum_{n=0}^{\infty} a_n z^n, \quad a_n \in \mathbb{K}$$

in the variable z. For $r \in [0, \infty[$ we define

$$\|f\|_r := \sum_{n=0}^{\infty} |a_n| r^n \in [0, \infty].$$

We write $\mathbb{K}[[z]]_r$ for the subset of all power series with $\|f\|_r < \infty$. Note that this implies that f converges uniformly to a function on the closed disc of radius r in $\mathbb{K}$.

Proposition 2.2.1. *Let $\mathcal{A}$ be a unital Banach algebra.*

(1) *If $x \in \mathcal{A}$ and $f \in \mathbb{K}[[z]]_r$ for some $r \geq \|x\|$, then $f(x) := \sum_{n=0}^{\infty} a_n x^n$ converges absolutely with*

$$\|f(x)\| \leq \|f\|_r.$$

For two power series $f(z) = \sum_n a_n z^n$ and $g(z) = \sum_n b_n z^n$ with $\|f\|_r, \|g\|_r < \infty$, we also have the product formula

$$(f \cdot g)(x) = f(x)g(x), \qquad \text{where} \qquad (f \cdot g)(z) := \sum_{n=0}^{\infty} \Big(\sum_{k=0}^{n} a_k b_{n-k} \Big) z^n \tag{2.4}$$

is the power series defined by the Cauchy product of f and g.

(2) *Suppose that $f(z) = \sum_{n=0}^{\infty} a_n z^n \in \mathbb{K}[[z]]_r$ and $g(z) = \sum_{n=1}^{\infty} b_n z^n \in z\mathbb{K}[[z]]$ satisfies $\|g\|_s < r$. We define the power series $f \circ g$ by formal composition:*

$$(f \circ g)(z) := \sum_n c_n z^n, \quad c_n = \sum_{k=0}^{n} a_k \sum_{i_1 + \cdots + i_k = n} b_{i_1} \cdots b_{i_k}.$$

Then $\|f \circ g\|_s \leq \|f\|_r$, and for any $x \in \mathcal{A}$ with $\|x\| \leq s$ the element $g(x)$ exists with $\|g(x)\| < r$, and we have the Composition Formula:

$$f(g(x)) = (f \circ g)(x). \tag{2.5}$$

Proof. (1) The convergence of $f(x)$ follows immediately from

$$\sum_n \|a_n x^n\| \leq \sum_n |a_n| \|x\|^n \leq \sum_n |a_n| r^n = \|f\|_r$$

and the Domination Test for absolutely converging series in a Banach space. We also obtain immediately the estimate $\|f(x)\| \leq \|f\|_r$.

If $\|f\|_r, \|g\|_r < \infty$, then (2.4) follows from the Cauchy Product Formula (Exercise 2.1.2) because the series $f(x)$ and $g(x)$ converge absolutely.

(2) To see that $\|f \circ g\|_s < \infty$, we calculate

$$\begin{aligned}\sum_n |c_n| s^n &\leq \sum_n \sum_{k=0}^n |a_k| \sum_{i_1+\cdots+i_k=n} |b_{i_1}| \cdots |b_{i_k}| s^n \\ &\leq \sum_{k=0}^\infty |a_k| \sum_n \sum_{i_1+\cdots+i_k=n} |b_{i_1}| \cdots |b_{i_k}| s^n = \sum_{k=0}^\infty |a_k| \|g\|_s^n \\ &\leq \sum_{k=0}^\infty |a_k| r^n = \|f\|_r.\end{aligned}$$

For $\|x\| \leq s$ we obtain from (1) the relation $\|g(x)\| \leq \|g\|_s$, so that

$$f(g(x)) = \sum_{n=0}^\infty a_n g(x)^n$$

is defined. Applying the Product Formula to the powers of g, we further obtain $g(x)^n = (g^n)(x)$, so that the polynomials $f_N(z) := \sum_{n=0}^N a_n z^n$ satisfy

$$f_N(g(x)) = \sum_{n=0}^N a_n g(x)^n = (f_N \circ g)(x).$$

Next we observe that

$$\|f \circ g - f_N \circ g\|_s = \|(f - f_N) \circ g\|_s \leq \|f - f_N\|_r \to 0,$$

so that

$$f_N(g(x)) = (f_N \circ g)(x) \to (f \circ g)(x).$$

Since we also have $f_N(g(x)) \to f(g(x))$ by definition, the Composition Formula is proved. □

Next we apply the preceding results to the logarithm series. Since this series has the radius of convergence 1, it defines a smooth function $\mathcal{A}^\times \supseteq B_1(\mathbf{1}) \to \mathcal{A}$, and we shall see that it provides an inverse of the exponential function.

Lemma 2.2.2. *The series* $\log(\mathbf{1} + x) := \sum_{k=1}^\infty (-1)^{k+1} \frac{x^k}{k}$ *converges for* $x \in \mathcal{A}$ *with* $\|x\| < 1$ *and defines a smooth function*

$$\log\colon B_1(\mathbf{1}) \to \mathcal{A}.$$

For $\|x\| < 1$ *and* $y \in \mathcal{A}$ *with* $xy = yx$ *we have*

$$(\mathtt{d}\log)(\mathbf{1} + x)y = (\mathbf{1} + x)^{-1} y.$$

Proof. The convergence follows from

$$\sum_{k=1}^{\infty}(-1)^{k+1}\frac{r^k}{k} = \log(1+r) < \infty$$

for $|r| < 1$, so that the smoothness follows from Theorem B.3.7.

If x and y commute, then the formula for the derivative in Theorem B.3.7 leads to

$$(\mathtt{d}\log)(\mathbf{1}+x)y = \sum_{k=1}^{\infty}(-1)^{k+1}x^{k-1}y = (\mathbf{1}+x)^{-1}y$$

(see the proof of Proposition 1.1.9). □

Proposition 2.2.3. (a) *For $x \in \mathcal{A}$ with $\|x\| < \log 2$ we have*

$$\log(\exp x) = x.$$

(b) *For $a \in \mathcal{A}^\times$ with $\|a - \mathbf{1}\| < 1$ we have $\exp(\log a) = a$.*

Proof. (a) We apply Proposition 2.2.1 with $g(z) = \exp(z) - 1 \in \bigcap_{s>0}\mathbb{K}[[z]]_s$ and $\log(\mathbf{1}+z) \in \mathbb{K}[[z]]_r$ for any $r < 1$. For $s < \log 2$ we then have $\|g\|_s \leq e^s - 1 < 1$. We thus obtain $\log(\exp x) = \log(\mathbf{1} + (\exp x - \mathbf{1})) = x$ for $\|x\| < \log 2$ from the formal relation $(\log\circ\exp)(z) = z$, which follows from the corresponding relation for the associated function on the real interval $]-\log 2, \log 2[$.

(b) Next we apply Proposition 2.2.1 with $f(z) = \exp(z)$ and $g(z) = \log(1+z)$ to obtain $\exp(\log a) = a$ for $\|a - \mathbf{1}| < 1$. □

Product and Commutator Formula

We have seen in Lemma 2.1.1 that the exponential image of a sum $x + y$ can be computed easily if x and y commute. In this case we also have for the commutator $[x, y] := xy - yx = 0$ the formula $\exp[x, y] = \mathbf{1}$. The following proposition gives a formula for $\exp(x + y)$ and $\exp([x, y])$ in the general case. The most natural way to obtain this formula, is by deriving it from the following lemma.

Lemma 2.2.4. *Let $\varepsilon > 0$ and $\gamma\colon [0, \varepsilon] \to \mathcal{A}$ be a continuous curve with $\gamma(0) = \mathbf{1}$. If $\gamma'(0)$ exists, then*

$$\lim_{n\to\infty}\gamma\Big(\frac{1}{n}\Big)^n = e^{\gamma'(0)}.$$

If, in addition, $\gamma'(0) = 0$, γ is C^1, and $\gamma''(0)$ exists, then

$$\lim_{n\to\infty}\gamma\Big(\frac{1}{n}\Big)^{n^2} = e^{\frac{\gamma''(0)}{2}}.$$

Proof. Since exp maps a neighborhood of 0 diffeomorphically onto a neighborhood of $\mathbf{1}$ and $\mathtt{d}\exp(0) = \mathrm{id}_{\mathcal{A}}$, we can, after possibly shrinking ε, write $\gamma(t) = e^{\beta(t)}$ with $\beta(0) = 0$ and $\beta'(0) = \gamma'(0)$. Then

$$\lim_{n\to\infty}\gamma\big(\frac{1}{n}\big)^n = \lim_{n\to\infty}e^{n\beta(\frac{1}{n})} \to e^{\gamma'(0)}$$

follows from $\beta(\frac{1}{n})n \to \beta'(0)$ and the continuity of exp.

If, in addition, $\gamma'(0) = 0$, γ is C^1 and $y := \gamma''(0)$ exists, then we put $\delta(t) := \gamma(\sqrt{t})$. Then the Fundamental Theorem of Calculus implies that

$$\delta(t) = \int_0^{\sqrt{t}} \gamma'(\tau)\, d\tau = \int_0^t \frac{1}{2\sqrt{s}} \gamma'(\sqrt{s})\, ds,$$

and since the continuous integrand converges to $y/2$ for $s \to 0$, we obtain for its mean value $\lim_{t\to 0} \delta(t)/t = y/2$. This shows that $\delta'(0) = y/2$ exists. From above we now obtain

$$\lim_{n\to\infty} \gamma\Big(\frac{1}{n}\Big)^{n^2} = \lim_{n\to\infty} \delta\Big(\frac{1}{n^2}\Big)^{n^2} = e^{y/2}. \qquad \square$$

Example 2.2.5. Applying the preceding lemma to the smooth curve $\gamma(t) := \mathbf{1} + tx$, we obtain the well-known formula

$$\lim_{n\to\infty} \Big(\mathbf{1} + \frac{x}{n}\Big)^n = e^x$$

for the exponential function.

If g, h are elements of a group G, then $(g, h) := ghg^{-1}h^{-1}$ is called their *commutator*. On the other hand, we call for two element $a, b \in \mathcal{A}$ the expression

$$[a, b] := ab - ba$$

their *commutator bracket*.

Proposition 2.2.6. *For $x, y \in \mathcal{A}$, the following assertions hold:*

(i) $\lim_{k\to\infty} \big(e^{\frac{1}{k}x} e^{\frac{1}{k}y}\big)^k = e^{x+y}$ (Trotter Product Formula).

(ii) $\lim_{k\to\infty} \big(e^{\frac{1}{k}x} e^{\frac{1}{k}y} e^{-\frac{1}{k}x} e^{-\frac{1}{k}y}\big)^{k^2} = e^{xy-yx}$ (Commutator Formula).

Proof. To obtain the product formula, we consider the smooth curve $\gamma(t) := e^{tx}e^{ty}$ with $\gamma(0) = \mathbf{1}$ and $\gamma'(0) = x + y$ (Product Rule). The assertion now follows from Lemma 2.2.4.

For the commutator formula, we consider the smooth curve $\gamma(t) := e^{tx}e^{ty}e^{-tx}e^{-ty}$. Then $e^{tx} = \mathbf{1} + tx + \frac{t^2}{2}x^2 + O(t^3)$ leads to

$$\begin{aligned}\gamma(t) &= \big(\mathbf{1} + tx + \frac{t^2}{2}x^2 + O(t^3)\big)\big(\mathbf{1} + ty + \frac{t^2}{2}y^2 + O(t^3)\big)\\ &\quad\cdot \big(\mathbf{1} - tx + \frac{t^2}{2}x^2 + O(t^3)\big)\big(\mathbf{1} - ty + \frac{t^2}{2}y^2 + O(t^3)\big)\\ &= \mathbf{1} + t(x + y - x - y) + t^2(x^2 + y^2 + xy - x^2 - xy - yx - y^2 + xy) + O(t^3)\\ &= \mathbf{1} + t^2(xy - yx) + O(t^3).\end{aligned}$$

This implies that $\gamma'(0) = 0$ and $\gamma''(0) = 2(xy - yx)$. Therefore the Commutator Formula follows from the second part of Lemma 2.2.4. $\square$

2.3 The Baker–Campbell–Dynkin–Hausdorff Formula

In this section we derive a formula which expresses the product $\exp x \exp y$ of two sufficiently small elements as the exponential image $\exp(x*y)$ of an element $x*y$ which can be described in terms of iterated commutator brackets. This implies in particular that the group multiplication in a small $\mathbf{1}$-neighborhood of $\mathcal{A}^\times$ is completely determined by the commutator bracket. To obtain these results, we express $\log(\exp x \exp y)$ as a power series $x * y$ in two variables. The (local) multiplication $*$ is called the *Baker–Campbell–Dynkin–Hausdorff multiplication* and the identity

$$\log(\exp x \exp y) = x * y$$

the *Baker–Campbell–Dynkin–Hausdorff formula* (BCDH). To make $x*y$ more explicit, we need some preparation. We start with the *adjoint representation of* $\mathcal{A}^\times$. This is the group homomorphism

$$\mathrm{Ad}\colon \mathcal{A}^\times \to \mathrm{Aut}(\mathcal{A}), \quad \mathrm{Ad}(g)x = gxg^{-1},$$

where $\mathrm{Aut}(\mathcal{A})$ stands for the group of algebra automorphisms of $\mathcal{A}$. For $x \in \mathcal{A}$, we further define a linear map representation

$$\mathrm{ad}(x)\colon \mathcal{A} \to \mathcal{A}, \quad \mathrm{ad}\, x(y) := [x, y] = xy - yx.$$

Lemma 2.3.1. *For each $x \in \mathcal{A}$,*

$$\mathrm{Ad}(\exp x) = \exp(\mathrm{ad}\, x). \tag{2.6}$$

Proof. We define the linear maps

$$\lambda_x\colon \mathcal{A} \to \mathcal{A}, \quad y \mapsto xy, \qquad \rho_x\colon \mathcal{A} \to \mathcal{A}, \quad y \mapsto yx.$$

Then $\lambda_x\rho_x = \rho_x\lambda_x$ and $\mathrm{ad}\, x = \lambda_x - \rho_x$, so that Lemma 2.1.1(ii) leads to

$$\mathrm{Ad}(\exp x)y = e^x y e^{-x} = e^{\lambda_x} e^{-\rho_x} y = e^{\lambda_x - \rho_x} y = e^{\mathrm{ad}\, x} y.$$

This proves (2.6). □

Proposition 2.3.2. *Let $x \in \mathcal{A}$ and $\lambda_{\exp x}(y) := (\exp x)y$ be the left multiplication by $\exp x$. Then*

$$\mathtt{d}\exp(x) = \lambda_{\exp x} \circ \frac{\mathbf{1} - e^{-\mathrm{ad}\, x}}{\mathrm{ad}\, x} \colon \mathcal{A} \to \mathcal{A},$$

where the fraction on the right means $\Phi(\mathrm{ad}\, x)$ for the entire function

$$\Phi(z) := \frac{1 - e^{-z}}{z} = \sum_{k=0}^{\infty} \frac{(-z)^k}{(k+1)!}.$$

Proof. Let $\alpha\colon [0,1] \to \mathcal{A}$ be a smooth curve. Then the map

$$\gamma\colon [0,1]^2 \to \mathcal{A}, \quad \gamma(t,s) := \exp(-s\alpha(t))\frac{d}{dt}\exp(s\alpha(t))$$

is C^1 in each argument and satisfies $\gamma(t,0) = 0$ for each t. We calculate

$$\begin{aligned}
\frac{\partial\gamma}{\partial s}(t,s) &= \exp(-s\alpha(t)) \cdot (-\alpha(t))\frac{d}{dt}\exp(s\alpha(t)) \\
&\quad + \exp(-s\alpha(t)) \cdot \frac{d}{dt}\Big(\alpha(t)\exp(s\alpha(t))\Big) \\
&= \exp(-s\alpha(t)) \cdot (-\alpha(t))\frac{d}{dt}\exp(s\alpha(t)) \\
&\quad + \exp(-s\alpha(t)) \cdot \Big(\alpha'(t)\exp(s\alpha(t)) + \alpha(t)\frac{d}{dt}\exp(s\alpha(t))\Big) \\
&= \mathrm{Ad}(\exp(-s\alpha(t)))\alpha'(t) = e^{-s\,\mathrm{ad}\,\alpha(t)}\alpha'(t).
\end{aligned}$$

Integration over $[0,1]$ with respect to s now leads to

$$\gamma(t,1) = \gamma(t,0) + \int_0^1 e^{-s\,\mathrm{ad}\,\alpha(t)}\alpha'(t)\,ds = \int_0^1 e^{-s\,\mathrm{ad}\,\alpha(t)}\,ds \cdot \alpha'(t).$$

Next we note that, for $x \in \mathcal{A}$,

$$\begin{aligned}
\int_0^1 e^{-s\,\mathrm{ad}\,x}\,ds &= \int_0^1 \sum_{k=0}^\infty \frac{(-\,\mathrm{ad}\,x)^k}{k!}s^k\,ds = \sum_{k=0}^\infty(-\,\mathrm{ad}\,x)^k\int_0^1 \frac{s^k}{k!}\,ds \\
&= \sum_{k=0}^\infty \frac{(-\,\mathrm{ad}\,x)^k}{(k+1)!} = \Phi(\mathrm{ad}\,x).
\end{aligned}$$

We thus obtain for $\alpha(t) = x + ty$ with $\alpha(0) = x$ and $\alpha'(0) = y$ the relation

$$\exp(-x)\mathtt{d}\exp(x)y = \gamma(0,1) = \int_0^1 e^{-s\,\mathrm{ad}\,x}y\,ds = \Phi(\mathrm{ad}\,x)y. \qquad \square$$

Remark 2.3.3. The explicit formula for the derivative of the exponential function can be written in many different ways. Here is another one which is sometimes convenient:

$$\begin{aligned}
\mathtt{d}\exp(x)y &= e^x\frac{\mathbf{1} - e^{-\,\mathrm{ad}\,x}}{\mathrm{ad}\,x}y = e^x\sum_{k=0}^\infty\frac{(-1)^k}{(k+1)!}(\mathrm{ad}\,x)^k y \\
&= e^x\int_0^1 e^{-s\,\mathrm{ad}\,x}y\,ds = e^x\int_0^1 e^{-sx}ye^{sx}\,ds = \int_0^1 e^{(1-s)x}ye^{sx}\,ds.
\end{aligned}$$

Lemma 2.3.4. *For*

$$\Phi(z) = \frac{1 - e^{-z}}{z} := \sum_{k=1}^\infty(-1)^{k-1}\frac{z^{k-1}}{k!}, \qquad z \in \mathbb{C}$$

and

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

we have

$$\Psi(e^z)\Phi(z) = 1 \quad \textit{for} \quad z \in \mathbb{C}, |z| < \log 2.$$

Proof. If $|z| < \log 2$, then $|e^z - 1| < 1$ and we obtain from $\log(e^z) = z$:

$$\Psi(e^z)\Phi(z) = \frac{e^z z}{e^z - 1} \frac{1 - e^{-z}}{z} = 1. \qquad \square$$

In view of the Composition Formula (2.5) (Proposition 2.2.1), the same identity as in Lemma 2.3.4 holds if we insert matrices $L \in \mathcal{L}(\mathcal{A})$ with $\|L\| < \log 2$ into the power series Φ and Ψ:

$$\Psi(\exp L)\Phi(L) = (\Psi \circ \exp)(L)\Phi(L) = ((\Psi \circ \exp) \cdot \Phi)(L) = \mathrm{id}_{\mathcal{A}}. \tag{2.7}$$

Here we use that $\|L\| < \log 2$ implies that all expressions are defined and in particular that $\|\exp L - \mathbf{1}\| < 1$, as a consequence of the estimate

$$\|\exp L - \mathbf{1}\| \leq e^{\|L\|} - 1. \tag{2.8}$$

The derivation of the BCDH formula follows a similar scheme as the proof of Proposition 2.3.2. Here we consider $x, y \in \mathcal{A}$ with $\|x\|, \|y\| < \log \sqrt{2}$. For $\|x\|, \|y\| < r$, the estimate (2.8) leads to

$$\begin{aligned} \|\exp x \exp y - \mathbf{1}\| &= \|(\exp x - \mathbf{1})(\exp y - \mathbf{1}) + (\exp y - \mathbf{1}) + (\exp x - \mathbf{1})\| \\ &\leq \|\exp x - \mathbf{1}\| \cdot \|\exp y - \mathbf{1}\| + \|\exp y - \mathbf{1}\| + \|\exp x - \mathbf{1}\| \\ &< (e^r - 1)^2 + 2(e^r - 1) = e^{2r} - 1. \end{aligned}$$

For $r < \log \sqrt{2} = \frac{1}{2} \log 2$ and $|t| \leq 1$, we obtain in particular

$$\|\exp x \exp ty - \mathbf{1}\| < e^{\log 2} - 1 = 1.$$

Therefore $\exp x \exp ty$ lies for $|t| \leq 1$ in the domain of the logarithm function (Lemma 2.2.2). We therefore define for $t \in [-1, 1]$:

$$F(t) = \log(\exp x \exp ty).$$

To estimate the norm of $F(t)$, we note that for $g := \exp x \exp ty$, $|t| \leq 1$, and $\|x\|, \|y\| < r$ we have

$$\begin{aligned} \|\log g\| &\leq \sum_{k=1}^{\infty} \frac{\|g - 1\|^k}{k} = -\log(1 - \|g - \mathbf{1}\|) \\ &< -\log(1 - (e^{2r} - 1)) = -\log(2 - e^{2r}). \end{aligned}$$

For $r := \frac{1}{2}\log(2 - \frac{\sqrt{2}}{2}) < \frac{\log 2}{2} = \log\sqrt{2}$ and $\|x\|, \|y\| < r$, this leads to

$$\|F(t)\| < -\log(2 - e^{2r}) = \log(\tfrac{2}{\sqrt{2}}) = \log(\sqrt{2}). \tag{2.9}$$

Next we calculate $F'(t)$ with the goal to obtain the BCDH formula as $F(1) = F(0) + \int_0^1 F'(t)\,dt$. For the derivative of the curve $t \mapsto \exp F(t)$, we get

$$\begin{aligned}(\mathtt{d}\exp)\big(F(t)\big)F'(t) &= \frac{d}{dt}\exp(F(t)) = \frac{d}{dt}\exp x \exp ty \\ &= (\exp x \exp ty)y = (\exp F(t))y.\end{aligned}$$

Using Proposition 2.3.2, we obtain

$$\begin{aligned} y &= \big(\exp F(t)\big)^{-1}(\mathtt{d}\exp)\big(F(t)\big)F'(t) \\ &= \frac{\mathbf{1} - e^{-\operatorname{ad} F(t)}}{\operatorname{ad} F(t)} F'(t) = \Phi(\operatorname{ad} F(t))F'(t). \end{aligned} \tag{2.10}$$

We claim that $\|\operatorname{ad}(F(t))\| < \log 2$. From $\|ab - ba\| \leq 2\|a\|\,\|b\|$ we derive

$$\|\operatorname{ad} a\| \leq 2\|a\| \quad \text{for} \quad a \in \mathcal{A}.$$

Therefore, by (2.9),

$$\|\operatorname{ad} F(t)\| \leq 2\|F(t)\| < 2\log(\sqrt{2}) = \log 2,$$

so that (2.10) and (2.7) lead to

$$F'(t) = \Phi(\operatorname{ad} F(t))^{-1} y = \Psi\big(\exp(\operatorname{ad} F(t))\big)y. \tag{2.11}$$

Proposition 2.3.5. *For $x, y \in \mathcal{A}$ with $\|x\|, \|y\| < \frac{1}{2}\log(2 - \frac{\sqrt{2}}{2})$, we have*

$$\log(\exp x \exp y) = x + \int_0^1 \Psi\big(\exp(\operatorname{ad} x)\exp(t \operatorname{ad} y)\big)y\,dt \in \mathcal{A},$$

for $\Psi(z) = \frac{z \log z}{z-1}$ as in Lemma 2.3.4.

Proof. Lemma 2.3.1 and the preceding remarks lead to

$$\begin{aligned}\exp(\operatorname{ad} F(t)) &= \operatorname{Ad}(\exp F(t)) = \operatorname{Ad}(\exp x \exp ty) = \operatorname{Ad}(\exp x)\operatorname{Ad}(\exp ty) \\ &= \exp(\operatorname{ad} x)\exp(\operatorname{ad} ty).\end{aligned}$$

With (2.11), this leads to

$$F'(t) = \Psi\big(\exp(\operatorname{ad} F(t))\big)y = \Psi\big(\exp(\operatorname{ad} x)\exp(\operatorname{ad} ty)\big)y.$$

Moreover, we have $F(0) = \log(\exp x) = x$. By integration we therefore obtain

$$\log(\exp x \exp y) = x + \int_0^1 \Psi\big(\exp(\operatorname{ad} x)\exp(t \operatorname{ad} y)\big)y\,dt. \qquad \square$$

Proposition 2.3.6. *For $x, y \in \mathcal{A}$ and $\|x\|, \|y\| < \frac{1}{2}\log(2 - \frac{\sqrt{2}}{2})$,*

$$\begin{aligned} x * y &:= \log(\exp x \exp y) \\ &= x + \\ &\sum_{\substack{k,m\geq 0 \\ p_i+q_i>0}} \frac{(-1)^k}{(k+1)(q_1+\ldots+q_k+1)} \frac{(\operatorname{ad} x)^{p_1}(\operatorname{ad} y)^{q_1}\ldots(\operatorname{ad} x)^{p_k}(\operatorname{ad} y)^{q_k}(\operatorname{ad} x)^m}{p_1!q_1!\ldots p_k!q_k!m!} y. \end{aligned}$$

Proof. We only have to rewrite the expression in Proposition 2.3.5:

$$\begin{aligned} &\int_0^1 \Psi\big(\exp(\operatorname{ad} x)\exp(\operatorname{ad} ty)\big) y\, dt \\ &= \int_0^1 \sum_{k=0}^{\infty} \frac{(-1)^k\big(\exp(\operatorname{ad} x)\exp(\operatorname{ad} ty) - \operatorname{id}\big)^k}{(k+1)} \big(\exp(\operatorname{ad} x)\exp(\operatorname{ad} ty)\big) y\, dt \\ &= \int_0^1 \sum_{\substack{k\geq 0 \\ p_i+q_i>0}} \frac{(-1)^k}{(k+1)} \frac{(\operatorname{ad} x)^{p_1}(\operatorname{ad} ty)^{q_1}\ldots(\operatorname{ad} x)^{p_k}(\operatorname{ad} ty)^{q_k}}{p_1!q_1!\ldots p_k!q_k!} \exp(\operatorname{ad} x) y\, dt \\ &= \sum_{\substack{k,m\geq 0 \\ p_i+q_i>0}} \frac{(-1)^k}{(k+1)} \frac{(\operatorname{ad} x)^{p_1}(\operatorname{ad} y)^{q_1}\ldots(\operatorname{ad} x)^{p_k}(\operatorname{ad} y)^{q_k}(\operatorname{ad} x)^m}{p_1!q_1!\ldots p_k!q_k!m!} y \int_0^1 t^{q_1+\ldots+q_k}\, dt \\ &= \sum_{\substack{k,m\geq 0 \\ p_i+q_i>0}} \frac{(-1)^k(\operatorname{ad} x)^{p_1}(\operatorname{ad} y)^{q_1}\ldots(\operatorname{ad} x)^{p_k}(\operatorname{ad} y)^{q_k}(\operatorname{ad} x)^m y}{(k+1)(q_1+\ldots+q_k+1)p_1!q_1!\ldots p_k!q_k!m!}. \end{aligned}$$

□

The series in Proposition 2.3.6 is called the *Hausdorff Series*. For practical purposes it often suffices to know the first terms of the Hausdorff Series:

Corollary 2.3.7. *For $x, y \in \mathcal{A}$ and $\|x\|, \|y\| < \frac{1}{2}\log(2 - \frac{\sqrt{2}}{2})$,*

$$x * y = x + y + \frac{1}{2}[x,y] + \frac{1}{12}[x,[x,y]] + \frac{1}{12}[y,[y,x]] + \ldots$$

Proof. One has to collect the summands in Proposition 2.3.6 corresponding to $p_1 + q_1 + \ldots + p_k + q_k + m \leq 2$. □

Chapter 3

Linear Lie Groups

In Section 3.1 we use the exponential function to associate to each closed subgroup $G \subseteq \mathcal{A}^\times$ a Banach–Lie algebra $\mathbf{L}(G)$, called the *Lie algebra of* G. We then show that the elements of $\mathbf{L}(G)$ are in one-to-one correspondence with the one-parameter groups of G and study some functorial properties of the assignment $\mathbf{L}\colon G \mapsto \mathbf{L}(G)$. Section 3.2 is devoted to some tools to calculate the Lie algebras of closed subgroups of $\mathcal{A}$.

3.1 Closed Subgroups of Banach Algebras

We call a subgroup $G \subseteq \mathcal{A}^\times$ of a unital Banach algebra $\mathcal{A}$ a *linear group*. In this section we shall use the exponential function to assign to each closed linear group G a vector space

$$\mathbf{L}(G) := \{x \in \mathcal{A}\colon \exp(\mathbb{R}x) \subseteq G\},$$

called the *Lie algebra of* G. This subspace carries a rich algebraic structure because for $x, y \in \mathbf{L}(G)$ the commutator $[x, y] = xy - yx$ is contained in $\mathbf{L}(G)$, so that $[\cdot,\cdot]$ defines a skew-symmetric bilinear operation on $\mathbf{L}(G)$. As a first step, we shall see how to calculate $\mathbf{L}(G)$ for concrete groups. Since the algebra $\mathcal{A} = M_n(\mathbb{R})$ of real $(n \times n)$-matrices also is a Banach algebra, all these results apply in particular to closed subgroups of $\mathrm{GL}_n(\mathbb{R})$.

The Lie Algebra of a Closed Linear Group

We start with the introduction of the concept of a Lie algebra.

Definition 3.1.1. (a) Let $\mathbb{K}$ be a field and L a $\mathbb{K}$-vector space. A bilinear map $[\cdot,\cdot]\colon L \times L \to L$ is called a *Lie bracket* if

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

(L2) $\big[x, [y, z]\big] = \big[[x, y], z\big] + \big[y, [x, z]\big]$ for $x, y, z \in L$ (Jacobi identity).[1]

[1] Carl Gustav Jacob Jacobi (1804–1851), mathematician in Berlin and Königsberg (Kaliningrad). He found his famous identity about 1830 in the context of Poisson brackets, which are related to Hamiltonian Mechanics and Symplectic Geometry.

A *Lie algebra* [2] (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* $\varphi\colon L_1 \to L_2$ of Lie algebras is a linear map with $\varphi([x,y]) = [\varphi(x), \varphi(y)]$ for $x, y \in L_1$. A Lie algebra is said to be *abelian* if $[x,y] = 0$ holds for all $x, y \in L$.

A *Banach–Lie algebra* is a Banach space L, endowed with a Lie algebra structure for which the bracket $[\cdot,\cdot]$ is continuous, i.e., there exists a $C > 0$ with

$$\|[x,y]\| \leq C\|x\| \cdot \|y\| \quad \text{for} \quad x, y \in L.$$

The following lemma shows that each associative algebra also carries a natural Lie algebra structure.

Lemma 3.1.2. *Each associative algebra $\mathcal{A}$ is a Lie algebra $\mathcal{A}_L$ with respect to the commutator bracket*

$$[a,b] := ab - ba.$$

If $\mathcal{A}$ is Banach algebra, then $\mathcal{A}_L$ is a Banach–Lie algebra with

$$\|[a,b]\| \leq 2\|a\|\|b\| \quad \textit{for} \quad a, b \in \mathcal{A}.$$

Proof. (L1) is obvious. For (L2) we calculate

$$[a, bc] = abc - bca = (ab - ba)c + b(ac - ca) = [a,b]c + b[a,c],$$

and this implies

$$[a,[b,c]] = [a,b]c + b[a,c] - [a,c]b - c[a,b] = [[a,b],c] + [b,[a,c]].$$

If, in addition, $\mathcal{A}$ is a Banach algebra, then the norm on $\mathcal{A}$ is submultiplicative, and this leads to

$$\|[x,y]\| = \|xy - yx\| \leq \|x\|\|y\| + \|y\|\|x\| = 2\|x\|\|y\|. \qquad \square$$

Definition 3.1.3. Let $\mathcal{A}$ be a unital Banach algebra. A subgroup $G \subseteq \mathcal{A}^\times$ is called a *linear group*. For each subgroup $G \subseteq \mathcal{A}$, we define the set

$$\mathbf{L}(G) := \{x \in \mathcal{A}\colon\ \exp(\mathbb{R}x) \subseteq G\}$$

and observe that $\mathbb{R}\,\mathbf{L}(G) \subseteq \mathbf{L}(G)$ follows immediately from the definition.

The next proposition assigns a Lie algebra to each closed linear group.

Proposition 3.1.4. *If $G \subseteq \mathcal{A}^\times$ is a closed subgroup, then $\mathbf{L}(G)$ is a closed real Lie subalgebra of $\mathcal{A}_L$.*

Proof. Let $x, y \in \mathbf{L}(G)$. For $k \in \mathbb{N}$ and $t \in \mathbb{R}$ we have $\exp \frac{t}{k}x, \exp \frac{t}{k}y \in G$ and with the Trotter Formula (Proposition 2.2.6), we get for all $t \in \mathbb{R}$:

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

[2]The notion of a Lie algebra was coined in the 1920s by Hermann Weyl.

because G is closed. Therefore $x + y \in \mathbf{L}(G)$.

Similarly we use the Commutator Formula to get

$$\exp t[x, y] = \lim_{k\to\infty} \left(\exp \frac{tx}{k} \exp \frac{y}{k} \exp -\frac{tx}{k} \exp -\frac{y}{k}\right)^{k^2} \in G,$$

hence $[x, y] \in \mathbf{L}(G)$.

That $\mathbf{L}(G)$ is closed follows from $\mathbf{L}(G) = \bigcap_{t\in\mathbb{R}} f_t^{-1}(G)$, for the continuous maps $f_t \colon \mathcal{A} \to \mathcal{A}^\times, x \mapsto e^{tx}$. □

Definition 3.1.5. In view of the preceding proposition, we obtain for each closed linear group G a map

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

which is called the *exponential function of* G.

The Banach–Lie algebra $\mathbf{L}(G)$ is called the *Lie algebra of* G. In particular,

$$\mathbf{L}(\mathcal{A}^\times) = \mathcal{A}_L.$$

Remark 3.1.6. If G is an abelian subgroup of $\mathcal{A}^\times$, then $\mathbf{L}(G)$ is also abelian.

Proposition 3.1.7. *Let $G \subseteq \mathcal{A}^\times$ be a subgroup. If* $\mathrm{Hom}(\mathbb{R}, G)$*, denotes the set of all continuous group homomorphisms* $(\mathbb{R}, +) \to G$*, then the map*

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

is a bijection.

Proof. For each $x \in \mathbf{L}(G)$, the map γ_x is a continuous group homomorphism (Theorem 2.1.5), and since $x = \gamma_x'(0)$, the map Γ is injective. To see that it is surjective, let $\gamma \colon \mathbb{R} \to G$ be a continuous group homomorphism and $\iota \colon G \to \mathcal{A}^\times$ the natural embedding. Then $\iota \circ \gamma \colon \mathbb{R} \to \mathcal{A}^\times$ is a continuous group homomorphism, so that there exists an $x \in \mathcal{A}$ with $\gamma(t) = \iota(\gamma(t)) = e^{tx}$ for all $t \in \mathbb{R}$ (Theorem 2.1.5). This implies that $x \in \mathbf{L}(G)$, and therefore that $\gamma_x = \gamma$. □

Examples 3.1.8. Let X be a Banach space.

(a) Then $\mathcal{L}(X)$ is a unital Banach algebra. We write $\mathrm{GL}(X) := \mathcal{L}(X)^\times$ for its unit group and $\mathfrak{gl}(X) := (\mathcal{L}(X), [\cdot,\cdot]) = \mathcal{L}(X)_L$ of its Lie algebra.

(b) Let $\widetilde{X} := X \times \mathbb{R}$. We consider the group homomorphism

$$\Phi \colon X \to \mathrm{GL}(\widetilde{X}), \quad x \mapsto \begin{pmatrix} \mathbf{1} & x \\ 0 & 1 \end{pmatrix}$$

and observe that Φ is an isomorphism of the topological group $(X, +)$ onto the closed linear group $\Phi(X)$.

The continuous one-parameter groups $\gamma \colon \mathbb{R} \to X$ are easily determined because $\gamma(nt) = n\gamma(t)$ for all $n \in \mathbb{Z}$, $t \in \mathbb{R}$, implies further $\gamma(q) = q\gamma(1)$ for all $q \in \mathbb{Q}$ and hence, by continuity, $\gamma(t) = t\gamma(1)$ for all $t \in \mathbb{R}$. Since $(X, +)$ is abelian, the Lie bracket on the Lie algebra

$$\mathbf{L}(\Phi(X)) = \left\{ \begin{pmatrix} \mathbf{0} & x \\ 0 & 0 \end{pmatrix} : x \in X \right\}$$

vanishes.

Definition 3.1.9. A *linear Lie group* is a closed subgroup G of the unit group $\mathcal{A}^\times$ of a unital Banach algebra $\mathcal{A}$ for which the exponential function

$$\exp_G\colon \mathbf{L}(G) \to G$$

is a local homeomorphism in 0, i.e., it maps some open 0-neighborhood U in $\mathbf{L}(G)$ homeomorphically onto an open $\mathbf{1}$-neighborhood in G.

Functorial Properties of the Lie Algebra

So far we have assigned to each closed linear group G its Lie algebra $\mathbf{L}(G)$. We shall also see that this assignment can be "extended" to continuous homomorphisms between closed linear groups in the sense that we assign to each such homomorphism

$$\varphi\colon G_1 \to G_2$$

a homomorphism $\mathbf{L}(\varphi)\colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ of Lie algebras, and this assignment satisfies

$$\mathbf{L}(\mathrm{id}_G) = \mathrm{id}_{\mathbf{L}(G)} \quad \text{and} \quad \mathbf{L}(\varphi_2 \circ \varphi_1) = \mathbf{L}(\varphi_2) \circ \mathbf{L}(\varphi_1)$$

for a composition $\varphi_1 \circ \varphi_2$ of two continuous homomorphisms $\varphi_1\colon G_2 \to G_1$ and $\varphi_2\colon G_3 \to G_2$. In the language of category theory, this means that $\mathbf{L}$ defines a functor from the category of linear Lie groups (where the morphisms are the continuous group homomorphisms) to the category of real Banach–Lie algebras.

Proposition 3.1.10. *Let $\varphi\colon G_1 \to G_2$ be a continuous group homomorphism of closed linear groups. Then the derivative*

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

exists for each $x \in \mathbf{L}(G_1)$ and defines a homomorphism of Lie algebras $\mathbf{L}(\varphi)\colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ with

$$\exp_{G_2} \circ \mathbf{L}(\varphi) = \varphi \circ \exp_{G_1}, \tag{3.1}$$

i.e., the following diagram commutes

$$\begin{array}{ccc} G_1 & \xrightarrow{\varphi} & G_2 \\ \uparrow{\scriptstyle \exp_{G_1}} & & \uparrow{\scriptstyle \exp_{G_2}} \\ \mathbf{L}(G_1) & \xrightarrow{\mathbf{L}(\varphi)} & \mathbf{L}(G_2). \end{array}$$

Then $\mathbf{L}(\varphi)$ is the uniquely determined linear map satisfying (3.1).

If, in addition, G_2 is a linear Lie group, then $\mathbf{L}(\varphi)$ is continuous.

Proof. For $x \in \mathbf{L}(G_1)$ we consider the homomorphism $\gamma_x \in \mathrm{Hom}(\mathbb{R}, G_1)$ given by $\gamma_x(t) = e^{tx}$. According to Proposition 3.1.7, we have

$$\varphi \circ \gamma_x(t) = \exp_{G_2}(ty)$$

for some $y \in \mathbf{L}(G_2)$, because $\varphi \circ \gamma_x \colon \mathbb{R} \to G_2$ is a continuous group homomorphism. Then clearly $y = (\varphi \circ \gamma_x)'(0) = \mathbf{L}(\varphi)x$. For $t = 1$ we obtain in particular

$$\exp_{G_2}(\mathbf{L}(\varphi)x) = \varphi(\exp_{G_1}(x)),$$

which is (3.1).

Conversely, every linear map $\psi \colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ with

$$\exp_{G_2} \circ \psi = \varphi \circ \exp_{G_1}$$

satisfies

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

and therefore

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

Next we show that $\mathbf{L}(\varphi)$ is a homomorphism of Lie algebras. From the definition of $\mathbf{L}(\varphi)$ we immediately get for $x \in \mathbf{L}(G_1)$:

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

which leads to $\mathbf{L}(\varphi)(tx) = t\,\mathbf{L}(\varphi)(x)$.

Since φ is continuous, the Trotter Formula implies that

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

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

We likewise obtain with the Commutator Formula

$$\varphi(\exp_{G_1}[x, y]) = \exp_{G_2}[\mathbf{L}(\varphi)(x), \mathbf{L}(\varphi)(y)]$$

and thus $\mathbf{L}(\varphi)([x, y]) = [\mathbf{L}(\varphi)(x), \mathbf{L}(\varphi)(y)]$.

If, in addition, H is a linear Lie group, then $\exp_H$ is a local homeomorphism in 0, so that the relation $\varphi \circ \exp_G = \exp_H \circ \mathbf{L}(\varphi)$ implies that $\mathbf{L}(\varphi)$ is continuous on some 0-neighborhood, and since it is a linear map, it is continuous (cf. Exercise 1.1.3). $\square$

Corollary 3.1.11. *If $\varphi_1 : G_1 \to G_2$ and $\varphi_2 \colon G_2 \to G_3$ are continuous homomorphisms of linear Lie groups, then*

$$\mathbf{L}(\varphi_2 \circ \varphi_1) = \mathbf{L}(\varphi_2) \circ \mathbf{L}(\varphi_1).$$

Moreover, $\mathbf{L}(\mathrm{id}_G) = \mathrm{id}_{\mathbf{L}(G)}$.

Proof. We have the relations

$$\varphi_1 \circ \exp_{G_1} = \exp_{G_2} \circ \mathbf{L}(\varphi_1) \quad \text{and} \quad \varphi_2 \circ \exp_{G_2} = \exp_{G_3} \circ \mathbf{L}(\varphi_2),$$

which immediately lead to

$$(\varphi_2 \circ \varphi_1) \circ \exp_{G_1} = \varphi_2 \circ \exp_{G_2} \circ \mathbf{L}(\varphi_1) = \exp_{G_3} \circ (\mathbf{L}(\varphi_2) \circ \mathbf{L}(\varphi_1)),$$

and the uniqueness assertion of Proposition 3.1.10 implies that

$$\mathbf{L}(\varphi_2 \circ \varphi_1) = \mathbf{L}(\varphi_2) \circ \mathbf{L}(\varphi_1).$$

Clearly $\mathrm{id}_{\mathbf{L}(G)}$ is a linear map satisfying $\exp_G \circ \mathrm{id}_{\mathbf{L}(G)} = \mathrm{id}_G \circ \exp_G$, so that the uniqueness assertion of Proposition 3.1.10 implies $\mathbf{L}(\mathrm{id}_G) = \mathrm{id}_{\mathbf{L}(G)}$. □

Corollary 3.1.12. *If $\varphi\colon G_1 \to G_2$ is an isomorphism of linear Lie groups, then $\mathbf{L}(\varphi)$ is an isomorphism of Banach–Lie algebras.*

Proof. Since φ is an isomorphism of linear Lie groups, it is bijective and $\psi := \varphi^{-1}$ also is a continuous homomorphism. We then obtain with Corollary 3.1.11 the relations $\mathrm{id}_{\mathbf{L}(G_2)} = \mathbf{L}(\mathrm{id}_{G_2}) = \mathbf{L}(\varphi \circ \psi) = \mathbf{L}(\varphi) \circ \mathbf{L}(\psi)$ and likewise $\mathrm{id}_{\mathbf{L}(G_1)} = \mathbf{L}(\psi) \circ \mathbf{L}(\varphi)$. Hence $\mathbf{L}(\varphi)$ is an isomorphism with $\mathbf{L}(\varphi)^{-1} = \mathbf{L}(\psi)$. □

Definition 3.1.13. If V is a vector space and G a group, then a homomorphism $\pi\colon G \to \mathrm{GL}(V)$ is called a *representation of G on V*. If $\mathfrak{g}$ is a Lie algebra, then a homomorphism of Lie algebras $\pi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is called a *representation of $\mathfrak{g}$ on V*.

As a consequence of Proposition 3.1.10, we obtain

Corollary 3.1.14. *If $\pi\colon G \to \mathrm{GL}(V)$ is a continuous representation of the closed linear group G on the Banach space V, then $\mathbf{L}(\pi)\colon \mathbf{L}(G) \to \mathfrak{gl}(V)$ is a representation of the Lie algebra $\mathbf{L}(G)$.*

Definition 3.1.15. The representation $\mathbf{L}(\pi)$ obtained in Corollary 3.1.14 from the group representation π is called the *derived representation*. This is motivated by the fact that for each $x \in \mathbf{L}(G)$ we have

$$\mathbf{L}(\pi)x = \frac{d}{dt}\Big|_{t=0} e^{t\,\mathbf{L}(\pi)x} = \frac{d}{dt}\Big|_{t=0} \pi(\exp tx).$$

The Adjoint Representation

Let $G \subseteq \mathcal{A}^\times$ be a linear Lie group and $\mathbf{L}(G) \subseteq \mathcal{A}$ the corresponding Lie algebra. For $g \in G$ we define the conjugation automorphism $c_g \in \mathrm{Aut}(G)$ by $c_g(x) := gxg^{-1}$. Then

$$\begin{aligned}\mathbf{L}(c_g)(x) &= \frac{d}{dt}\Big|_{t=0} c_g(\exp tx) = \frac{d}{dt}\Big|_{t=0} g(\exp tx)g^{-1} \\ &= \frac{d}{dt}\Big|_{t=0} \exp(tgxg^{-1}) = gxg^{-1}\end{aligned}$$

(Lemma 2.1.1), and therefore $\mathbf{L}(c_g) = c_g|_{\mathbf{L}(G)}$. We define the *adjoint representation of G on $\mathbf{L}(G)$* by

$$\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathbf{L}(G)), \quad \mathrm{Ad}(g)(x) := \mathbf{L}(c_g)x = gxg^{-1}.$$

(That this is a representation follows immediately from the explicit formula).

For each $x \in \mathbf{L}(G)$, the map $G \to \mathbf{L}(G), g \mapsto \mathrm{Ad}(g)(x) = gxg^{-1}$ is continuous and each $\mathrm{Ad}(g)$ is an automorphism of the Lie algebra $\mathbf{L}(G)$. Therefore Ad is a continuous homomorphism from the closed linear group G to the closed linear group $\mathrm{Aut}(\mathbf{L}(G)) \subseteq \mathrm{GL}(\mathbf{L}(G))$. The derived representation

$$\mathbf{L}(\mathrm{Ad})\colon \mathbf{L}(G) \to \mathfrak{gl}(\mathbf{L}(G))$$

is a representation of $\mathbf{L}(G)$ on $\mathbf{L}(G)$. We further define for $x \in \mathbf{L}(G)$ a linear map

$$\mathrm{ad}(x)\colon \mathbf{L}(G) \to \mathbf{L}(G), \quad \mathrm{ad}\, x(y) := [x, y].$$

Lemma 3.1.16. $\mathbf{L}(\mathrm{Ad}) = \mathrm{ad}$.

Proof. In view of Proposition 3.1.10, this is an immediate consequence of the relation $\mathrm{Ad}(\exp x) = e^{\mathrm{ad}\, x}$ (Lemma 2.3.1). □

Exercises for Section 3.1

Exercise 3.1.1. If $(G_j)_{j\in J}$ is a family of subgroups of $\mathcal{A}^\times$, then

$$\mathbf{L}\Big(\bigcap_{j\in J} G_j\Big) = \bigcap_{j\in J} \mathbf{L}(G_j).$$

Exercise 3.1.2. Let $G := \mathrm{GL}_n(\mathbb{K})$ and $V := P_k(\mathbb{K}^n)$ be the space of homogeneous polynomials of degree k in $x_1, \ldots, x_n$, considered as functions $\mathbb{K}^n \to \mathbb{K}$. Show that:

(1) $\dim V = \binom{k+n-1}{n-1}$.

(2) We obtain a continuous representation $\rho\colon G \to \mathrm{GL}(V)$ of G on V by $(\rho(g)f)(x) := f(g^{-1}x)$.

(3) For the elementary matrix $E_{ij} = (\delta_{ij})$ we have $\mathbf{L}(\rho)(E_{ij}) = -x_j \frac{\partial}{\partial x_i}$. Hint: $(\mathbf{1} + tE_{ij})^{-1} = \mathbf{1} - tE_{ij}$.

Exercise 3.1.3. If $X \in \mathrm{End}(V)$ is nilpotent, then $\mathrm{ad}\, X \in \mathrm{End}(\mathrm{End}(V))$ is also nilpotent. Hint: $\mathrm{ad}\, X = L_X - R_X$ and both summands commute.

Exercise 3.1.4. If $(V, \cdot)$ is an associative algebra, then we have $\mathrm{Aut}(V, \cdot) \subseteq \mathrm{Aut}(V, [\cdot, \cdot])$.

Exercise 3.1.5. Let V and W be vector spaces and $q : V \times V \to W$ a skew-symmetric bilinear map. Then

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

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

Exercise 3.1.6. Let $\mathfrak{g}$ be a Lie algebra with $\big[x,[y,z]\big] = 0$ for $x, y, z \in \mathfrak{g}$. Then

$$x * y := x + y + \frac{1}{2}[x, y]$$

defines a group structure on $\mathfrak{g}$. An example for such a Lie algebra is the three-dimensional *Heisenberg algebra*

$$\mathfrak{g} = \left\{ \begin{pmatrix} 0 & p & z \\ 0 & 0 & q \\ 0 & 0 & 0 \end{pmatrix} : p, q, z \in \mathbb{K} \right\}.$$

Exercise 3.1.7. Show that every Banach–Lie algebra L carries a norm $\|\cdot\|$ defining the topology for which

$$\|[x, y]\| \leq \|x\|\|y\| \quad \text{for} \quad x, y \in L.$$

Exercise 3.1.8. Let $G \subseteq \mathcal{A}^\times$ be a closed subgroup. We call $v \in \mathcal{A}$ *a tangent vector of* G *in* g if there exists a curve $\gamma\colon [0,1] \to G$ with $\gamma(0) = g$ for which $\gamma'(0) = v$ exists. We write $T_g(G)$ for the set of tangent vectors of G in g. Show that

(i) The set $T_{\mathbf{1}}(G)$ of tangent vectors of G in $\mathbf{1}$ coincides with the Lie algebra $\mathbf{L}(G)$. In particular it is a closed subspace.

(ii) $T_g(G) = g\,\mathbf{L}(G)$ for every $g \in G$.

Exercise 3.1.9. If $\mathcal{A}$ is a unital Banach algebra, then we endow the vector space $T\mathcal{A} := \mathcal{A} \oplus \mathcal{A}$ with the norm $\|(a, b)\| := \|a\| + \|b\|$ and the multiplication

$$(a, b)(a', b') := (aa', ab' + ba').$$

Show that

(1) $T\mathcal{A}$ is a unital Banach algebra with identity $(\mathbf{1}, 0)$. It is called the *tangent algebra of* $\mathcal{A}$.

(2) For $\varepsilon := (0, 1)$, each element of $T\mathcal{A}$ can be written in a unique fashion as $(a, b) = a + b\varepsilon$ and the multiplication satisfies

$$(a + b\varepsilon)(a' + b'\varepsilon) = aa' + (ab' + ba')\varepsilon.$$

In particular, $\varepsilon^2 = 0$.

(3) $(T\mathcal{A})^\times = \mathcal{A}^\times \times \mathcal{A}$.

(4) If $G \subseteq \mathcal{A}^\times$ is a closed subgroup, then its tangent group $TG := G \cdot (\mathbf{1} + \varepsilon\,\mathbf{L}(G))$ is a closed subgroup of $(T\mathcal{A})^\times$ and

$$\mathbf{L}(G) = \{x \in \mathcal{A}\colon \mathbf{1} + \varepsilon x \in T(G)\}.$$

Exercise 3.1.10. (a) For each closed subgroup $G \subseteq \mathcal{A}^\times$, the map $\mathrm{Ad} : G \to \mathrm{Aut}(\mathbf{L}(G))$ is a group homomorphism (called the *adjoint representation of* G).

(b) For each Lie algebra $\mathfrak{g}$, the map $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a homomorphism of Lie algebras (called the *adjoint representation of* $\mathfrak{g}$).

3.2 Calculating Lie Algebras of Linear Groups

In this section we shall see various techniques to determine the Lie algebra of a linear Lie group.

Example 3.2.1. Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Then the group $G := \mathrm{SL}_n(\mathbb{K}) = \det^{-1}(1) = \ker \det$ is a closed linear group. To determine its Lie algebra, we first claim that

$$\det(e^x) = e^{\operatorname{tr} x} \tag{3.2}$$

holds for $x \in M_n(\mathbb{K})$. To verify this claim, we consider

$$\det\colon M_n(\mathbb{K}) \cong (\mathbb{K}^n)^n \to \mathbb{K}$$

as an n-linear map, where each matrix x is considered as an n-tuple of its column vectors $x_1, \ldots, x_n$. Then Lemma B.2.2 implies that

$$\begin{aligned}&(\mathtt{d}\det)(\mathbf{1})(x) = (\mathtt{d}\det)(e_1, \ldots, e_n)(x_1, \ldots, x_n)\\ &= \det(x_1, e_2, \ldots, e_n) + \ldots + \det(e_1, \ldots, e_{n-1}, x_n) = x_{11} + \ldots + x_{nn} = \operatorname{tr} x.\end{aligned}$$

For the continuous group homomorphism $\det\colon \mathrm{GL}_n(\mathbb{K}) \to \mathbb{K}^\times = \mathrm{GL}_1(\mathbb{K})$ we therefore obtain

$$\mathbf{L}(\det) = \operatorname{tr}\colon \mathfrak{gl}_n(\mathbb{K}) \to \mathfrak{gl}_1(\mathbb{K}) \cong \mathbb{K},$$

so that (3.2) follows from $\det(e^x) = e^{\mathbf{L}(\det)x} = e^{\operatorname{tr} x}$. We conclude that

$$\begin{aligned}\mathfrak{sl}_n(\mathbb{K}) &:= \mathbf{L}(\mathrm{SL}_n(\mathbb{K})) = \{x \in M_n(\mathbb{K})\colon (\forall t \in \mathbb{R})\, 1 = \det(e^{tx}) = e^{t \operatorname{tr} x}\}\\ &= \{x \in M_n(\mathbb{K})\colon \operatorname{tr} x = 0\}.\end{aligned}$$

Lemma 3.2.2. *Let V and W be Banach spaces and $\beta\colon V \times V \to W$ a continuous bilinear map. For $(x, y) \in \mathfrak{gl}(V) \times \mathfrak{gl}(W)$, the following are equivalent:*

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

Proof. (a) $\Rightarrow$ (b): Taking the derivative in $t = 0$, the relation (a) leads to

$$y\beta(v, v') = \beta(xv, v') + \beta(v, xv'),$$

where we use the Product and the Chain Rule.

(b) $\Rightarrow$ (a): If (b) holds, then we obtain inductively

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

For the exponential series this leads with the general Cauchy Product Formula (Exercise 2.1.2) to

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

Since (b) also holds for the pair (tx, ty) for all $t \in \mathbb{R}$, this completes the proof. □

Proposition 3.2.3. *Let V and W be Banach space and $\beta\colon V \times V \to W$ a continuous bilinear map. For the group*

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

we then have

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

Proof. We only have to observe that $X \in \mathbf{L}(\mathrm{Aut}(V,\beta))$ is equivalent to the pair $(X, 0)$ satisfying condition (a) in Lemma 3.2.2. □

Example 3.2.4. (a) Let $B \in M_n(\mathbb{K})$, $\beta(v,w) = v^\top B w$, and

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

Then Proposition 3.2.3 implies that

$$\begin{aligned} \mathfrak{aut}(\mathbb{K}^n,\beta) := \mathbf{L}(\mathrm{Aut}(\mathbb{K}^n,\beta)) &= \{x \in \mathfrak{gl}_n(\mathbb{K})\colon (\forall v, v' \in V)\ \beta(xv, v') + \beta(v, xv') = 0\} \\ &= \{x \in \mathfrak{gl}_n(\mathbb{K})\colon (\forall v, v' \in V)\ v^\top x^\top B v' + v^\top B x v' = 0\} \\ &= \{x \in \mathfrak{gl}_n(\mathbb{K})\colon x^\top B + Bx = 0\}. \end{aligned}$$

In particular, we obtain for the *orthogonal group*

$$\mathrm{O}_n(\mathbb{K}) := \{g \in \mathrm{GL}_n(\mathbb{K})\colon g^\top = g^{-1}\}$$

the Lie algebra

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

Let $q := n - p$ and let $I_{p,q}$ denote the corresponding matrix

$$I_{p,q} = \begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_q \end{pmatrix} \in M_{p+q}(\mathbb{R}).$$

Then we obtain for the *indefinite orthogonal group*

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

the Lie algebra

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

and for the *symplectic group*

$$\mathrm{Sp}_{2n}(\mathbb{K}) := \{g \in \mathrm{GL}_{2n}(\mathbb{K})\colon g^\top B g = B\}, \quad B = \begin{pmatrix} 0 & -\mathbf{1}_n \\ \mathbf{1}_n & 0 \end{pmatrix},$$

we find

$$\mathfrak{sp}_{2n}(\mathbb{K}) := \mathbf{L}(\mathrm{Sp}_{2n}(\mathbb{K})) := \{x \in \mathfrak{gl}_{2n}(\mathbb{K})\colon x^\top B + Bx = 0\}.$$

(b) Applying Proposition 3.2.3 with $V = \mathbb{C}^n$ and $W = \mathbb{C}$, considered as real vector spaces, we also obtain for a hermitian form

$$\beta\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \quad (z, w) \mapsto w^* I_{p,q} z$$

and the corresponding automorphism group

$$\mathrm{U}_{p,q}(\mathbb{C}) := \mathrm{Aut}(\mathbb{C}^n, \beta)$$

the Lie algebra

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

In particular, we get

$$\mathfrak{u}_n(\mathbb{C}) := \mathbf{L}(\mathrm{U}_n(\mathbb{C})) = \{x \in \mathfrak{gl}_n(\mathbb{C})\colon x^* = -x\}.$$

(c) If $\mathcal{H}$ is a complex Hilbert space, then $\mathrm{U}(\mathcal{H})$ is a closed subgroup of $\mathrm{GL}(\mathcal{H})$, and we obtain for its Lie algebra

$$\mathfrak{u}(\mathcal{H}) := \mathbf{L}(\mathrm{U}(\mathcal{H})) = \{x \in \mathfrak{gl}(\mathcal{H})\colon x^* = -x\}.$$

To see that $\mathrm{U}(\mathcal{H})$ actually is a linear Lie group, let $U = -U = U^* \subseteq B(\mathcal{H})$ be an open symmetric $*$-invariant 0-neighborhood mapped by exp diffeomorphically onto an open subset of $\mathrm{GL}(\mathcal{H})$ (Proposition 2.1.6). Then, for $x \in U$, the relations

$$(e^x)^* = e^{x^*} \quad \text{and} \quad (e^x)^{-1} = e^{-x}$$

imply that

$$\exp(U) \cap \mathrm{U}(\mathcal{H}) = \{e^x\colon x^* = -x\} = \exp(U \cap \mathfrak{u}(\mathcal{H})).$$

Therefore $\mathrm{U}(\mathcal{H})$ is a linear Lie group.

Example 3.2.5. Let $\mathfrak{g}$ be a Banach Lie algebra and

$$\mathrm{Aut}(\mathfrak{g}) := \{g \in \mathrm{GL}(\mathfrak{g}) \colon (\forall x, y \in \mathfrak{g})\ g[x, y] = [gx, gy]\}.$$

Then $\mathrm{Aut}(\mathfrak{g})$ is a closed subgroup of $\mathrm{GL}(\mathfrak{g})$, in particular a linear group. To calculate the Lie algebra of G, we use Lemma 3.2.2 with $V = W = \mathfrak{g}$ and $\beta(x, y) = [x, y]$. Then we see that $D \in \mathfrak{aut}(\mathfrak{g}) := \mathbf{L}(\mathrm{Aut}(\mathfrak{g}))$ is equivalent to (D, D) satisfying the conditions in Lemma 3.2.2, and this leads to

$$\mathfrak{aut}(\mathfrak{g}) := \mathbf{L}(\mathrm{Aut}(\mathfrak{g})) = \{D \in \mathfrak{gl}(\mathfrak{g}) \colon (\forall x, y \in \mathfrak{g})\, D[x, y] = [Dx, y] + [x, Dy]\}$$

The elements of this Lie algebra are called *derivations of* $\mathfrak{g}$, and $\mathfrak{aut}(\mathfrak{g})$ is also denoted $\mathrm{der}(\mathfrak{g})$. Note that the condition on an endomorphism of $\mathfrak{g}$ to be a derivation resembles the Leibniz Rule (Product Rule).

Remark 3.2.6. If $\mathcal{A}$ is a complex unital Banach algebra, we call a closed linear group $G \subseteq \mathcal{A}^\times$ a *complex linear group* if $\mathbf{L}(G) \subseteq \mathcal{A}$ is a complex subspace, i.e., $i\,\mathbf{L}(G) \subseteq \mathbf{L}(G)$. Since Proposition 3.1.4 only ensures that $\mathbf{L}(G)$ is a real subspace, this requirement is not automatically satisfied.

If $\mathcal{H}$ is a complex Hilbert space, then the closed linear group $\mathrm{U}(\mathcal{H}) \subseteq \mathrm{GL}(\mathcal{H})$ is not a complex linear group because

$$i\mathfrak{u}(\mathcal{H}) = \mathrm{Herm}(\mathcal{H}) \not\subseteq \mathfrak{u}(\mathcal{H}).$$

This is due to the fact that the scalar product on $\mathcal{H}$ whose automorphism group is $\mathrm{U}(\mathcal{H})$, is not complex bilinear. For any complex bilinear form $\beta \colon V \times V \to \mathbb{C}$, the corresponding group $\mathrm{Aut}(V, \beta)$ is a complex linear group because

$$\mathfrak{aut}(V, \beta) = \{X \in \mathfrak{gl}(V) \colon (\forall v, w \in V)\, \beta(Xv, w) + \beta(v, Xw) = 0\}$$

is a complex subspace of $\mathfrak{gl}(V)$.

Exercises for Section 3.2

Exercise 3.2.1. Show that for $X = -X^* \in M_n(\mathbb{C})$ the matrix e^X is unitary and that the exponential function

$$\exp : \mathrm{Aherm}_n(\mathbb{C}) := \{X \in M_n(\mathbb{C}) \colon X^* = -X\} \to \mathrm{U}_n(\mathbb{C}), \quad X \mapsto e^X$$

is surjective.

Exercise 3.2.2. Show that for $X^\top = -X \in M_n(\mathbb{R})$ the matrix e^X is orthogonal and that the exponential function

$$\exp : \mathrm{Skew}_n(\mathbb{R}) := \{X \in M_n(\mathbb{R}) \colon X^\top = -X\} \to \mathrm{O}_n(\mathbb{R})$$

is not surjective. Can you determine which orthogonal matrices are contained in the image? Can you interprete the result geometrically in terns of the geometry of the flow $\mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n, (t, v) \mapsto e^{tX} v$.

Exercise 3.2.3. Show that a closed subgroup $G \subseteq \mathcal{A}^\times$ is a linear Lie group if and only if there exists an open 0-neighborhood $U \subseteq \mathbf{L}(G)$ for which $\exp(U)$ is a $\mathbf{1}$-neighborhood of G. Hint: Proposition 2.1.6.

Exercise 3.2.4. Let $\varphi\colon G \to H$ be a continuous homomorphism of linear Lie groups. Show that $\ker\varphi$ is a linear Lie group with Lie algebra

$$\mathbf{L}(\ker\varphi) = \ker \mathbf{L}(\varphi).$$

Exercise 3.2.5. Let $B \in \mathrm{GL}_n(\mathbb{K})$ and consider the bilinear form β on $\mathbb{K}^n$ defined by $\beta(v,w) := v^\top Bw$. Show that $\mathrm{Aut}(\mathbb{K}^n, \beta)$ is a linear Lie group. Hint: Show that $e^X \in \mathrm{Aut}(\mathbb{K}^n, \beta)$ is equivalent to $Be^X B^{-1} = e^{-X^\top}$.

Chapter 4

General Lie Groups

4.1 Manifolds and Lie Groups

We have already encountered linear Lie groups, which are certain closed subgroups of the unit group $\mathcal{A}^\times$ of a Banach algebra $\mathcal{A}$. In this section, we define Lie groups in the context of Banach manifolds. We explain how the Lie algebra and the corresponding Lie functor are defined and describe some basic properties.

Smooth manifolds

Contrary to submanifolds of some vector space, a differentiable manifold is described without specifying any surrounding space. In spite of the fact that one can show that each finite dimensional smooth manifold can be realized as a closed submanifold of some $\mathbb{R}^n$ (Whitney's Embedding Theorem), these embeddings are not canonical, and it is therefore much more natural to think of differentiable manifolds as spaces for which no embedding is specified. The concept of a differentiable manifold permits us to define a Lie group as a differentiable manifold for which the group operations are smooth maps. We shall verify below that this approach is compatible with what we have learned previously on linear Lie groups.

Definition 4.1.1. Let M be a Hausdorff space and E be a Banach space.

An *E-chart* of M is a pair (φ, U), where $U \subseteq M$ is an open subset and $\varphi\colon U \to E$ is a homeomorphism onto an open subset of E. If $E = \mathbb{R}^n$, then an E-chart is simply called an n-dimensional chart and we think of $\varphi(x) \in \mathbb{R}^n$ as an n-tuple of coordinates of the element $x \in M$.

Two E-charts (φ, U) and (ψ, V) are *compatible* if the map

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

and its inverse is smooth, i.e., if it is a diffeomorphism.

An *E-atlas on M* is a family $(\varphi_\alpha, U_\alpha)_{\alpha \in I}$ of E-charts with the following properties:

(M1) $M = \bigcup_{\alpha \in I} U_\alpha$.

(M2) For $\alpha, \beta \in I$, the charts $(\varphi_\alpha, U_\alpha)$ and (φ_β, U_β) are compatible.

We call an E-atlas $\mathcal{A}$ on M *maximal* if it contains all charts (ψ, V) compatible with all charts of $\mathcal{U}$. A maximal atlas on M is called a *differentiable structure*, and the pair $(M, \mathcal{U})$ is called a *differentiable/smooth manifold modelled on* E or a *smooth* E*-manifold*. Then $\dim M := \dim E$ is called the *dimension of* M. In the following we simply write M for $(M, \mathcal{U})$ and call the charts in the atlas $\mathcal{U}$ simply the charts of M.

Remark 4.1.2. (a) A given topological space M may carry different differentiable structures. Examples are the exotic differentiable structures on $\mathbb{R}^4$ (the only $\mathbb{R}^n$ carrying exotic differentiable structures) and the 7-sphere $\mathbb{S}^7$.

(b) Each atlas $\mathcal{U}$ on a Hausdorff space M is contained in a maximal atlas. One simply has to enlarge $\mathcal{U}$ by all the charts compatible with the charts in $\mathcal{U}$. It is easy to verify that one thus obtains an atlas. To specify the structure of a differentiable manifold on M, it therefore suffices to specify an atlas.

Example 4.1.3. (a) Each Banach space $M = E$ is a smooth manifold, where the differentiable structure is given by the atlas (E, id_E).

(b) Each open subset N of a smooth E-manifold M inherits a canonical E-manifold structure. Its charts are obtained by restricting a chart (φ, U) of M to $(\varphi|_{U\cap N}, U \cap N)$. In particular, all open subsets of Banach spaces carry natural manifold structures.

(c) If $M \subseteq \mathbb{R}^n$ is a k-dimensional submanifold, then M is a Hausdorff space with respect to the topology inherited from $\mathbb{R}^n$. To obtain an atlas on M, we consider for each $x \in M$ an open neighborhood V_x of x in $\mathbb{R}^n$ and a diffeomorphism $\psi\colon V_x \to W$ onto an open neighborhood W of 0 in $\mathbb{R}^n$ such that

$$\psi(V_x \cap M) = W \cap \mathbb{R}^k.$$

We define $U_x := M \cap V_x$ and $\varphi_x := \psi|_{U_x \cap M} \colon U_x \cap M \to \mathbb{R}^k$. Then $\mathcal{U} := (\varphi_x, U_x)_{x\in M}$ is an atlas of M. This requires some verification! (see Analysis II).

(d) An important example of a smooth manifold is the *sphere*

$$\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} \colon \|x\|_2 = 1\}.$$

For each $i \in \{1, \ldots, n+1\}$ we consider the open subsets

$$U_{i,\pm} := \{x \in \mathbb{S}^n \colon \pm x_i > 0\}.$$

It is clear that the open sets $U_{i,\pm}$ cover $\mathbb{S}^n$. On each $U_{i,\pm}$ we define a chart

$$\varphi_i \colon U_{i,\pm} \to \mathbb{R}^n, \quad (x_0, \ldots, x_n) \mapsto (x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

Then each φ_i is a homeomorphism of $U_{i,\pm}$ onto the open unit ball B in $\mathbb{R}^n$, and

$$\varphi_i^{-1} \colon B \to U_{i,\pm}, \quad y \mapsto \Big(y_0, \ldots, y_{i-1}, \pm\big(1 - \sum_{j\neq i} y_j^2\big)^{\frac{1}{2}}, y_{i+1}, \ldots, y_{n+1}\Big).$$

It is easy to verify that the charts $(\varphi_i, U_{i,\pm})$ form a smooth atlas on $\mathbb{S}^n$.

Definition 4.1.4. (a) Let M and N be differentiable manifolds. We call a continuous map $f\colon M \to N$ *smooth in* $p \in M$ if, for some chart (φ, U) of M with $p \in U$ and some chart (ψ, V) of N with $f(p) \in V$, the map

$$\psi \circ f \circ \varphi^{-1}\colon \varphi(f^{-1}(V)) \to \psi(V), \quad \varphi(x) \mapsto \psi(f(x)) \tag{4.1}$$

between open subsets of a Banach space is smooth in a neighborhood of $\varphi(p)$. Note that the assumption that f is continuous implies that $f^{-1}(U)$ is open in M, so that the set $\psi(f^{-1}(U))$ is open. We call a continuous map $f\colon M \to N$ *smooth* if it is smooth in each point of M.

(b) A smooth map $f\colon M \to N$ is called a *differentiable isomorphism* or a *diffeomorphism* if there exists a smooth map $g\colon N \to M$ with $g \circ f = \mathrm{id}_M$ and $f \circ g = \mathrm{id}_N$.

Remark 4.1.5. (a) If $f\colon M \to N$ and $g\colon N \to Q$ are continuous maps and $p \in M$ is such that f is smooth in p and g is smooth in $f(p)$, then the composition $g \circ f$ is smooth in p. In fact, for charts (φ, U), (ψ, V), resp., (η, W) of M, N, resp., Q, we have

$$\eta \circ (g \circ f) \circ \varphi^{-1} = (\eta \circ g \circ \psi^{-1}) \circ (\psi \circ f \circ \varphi^{-1}),$$

on its natural domain, which contains a neighborhood of $\varphi(p)$.

(b) From (a) it follows in particular that, if $f\colon M \to N$ is smooth in p and $(\widetilde{\varphi}, \widetilde{U})$ is any chart of M with $p \in \widetilde{U}$, then, for any chart $(\widetilde{\psi}, \widetilde{V})$ of N with $f(p) \in \widetilde{V}$, the map

$$\widetilde{\psi} \circ f \circ \widetilde{\varphi}^{-1}\colon \widetilde{\varphi}(f^{-1}(\widetilde{V})) \to \widetilde{\psi}(\widetilde{V})$$

is smooth.

(c) The map $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^3$ is smooth and invertible, but it is not a smooth isomorphism because $f^{-1}(x) = x^{1/3}$ is not differentiable in 0.

Definition 4.1.6. (Product manifolds) Let M, resp., N be differentiable manifolds with an E-atlas $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$, resp., and F-atlas $(V_\beta, \psi_\beta)_{\beta \in B}$. Then the topological product $M \times N$ is a Hausdorff space and we obtain the structure of a smooth $(E \times F)$-manifold on $M \times N$ by the atlas $(U_\alpha \times V_\beta, \varphi_\alpha \times \psi_\beta)_{(\alpha,\beta) \in A \times B}$, where

$$(\varphi_\alpha \times \psi_\beta)\colon U_\alpha \times V_\beta \to \varphi_\alpha(U_\alpha) \times \psi_\beta(V_\beta), \quad (x, y) \mapsto (\varphi_\alpha(x), \psi_\beta(y)).$$

Remark 4.1.7. If M and N are differentiable manifolds, then the product manifold $M \times N$ has the following properties:

(a) The projection maps $p_M\colon M \times N \to M$ and $p_N\colon M \times N \to N$ are smooth.

(b) For $x \in M$, the embedding

$$i_x\colon N \to M \times N, \quad y \mapsto (x, y)$$

is smooth and, for $y \in N$, the embedding

$$i^y\colon M \to M \times N, \quad x \mapsto (x, y)$$

is smooth.

(c) The diagonal embedding

$$\Delta_M\colon M \to M \times M, \quad x \mapsto (x, x)$$

is smooth.

The definition of a Lie group

There are two types of additional structures on groups. The first level consists of a topological structure compatible with the group structure, which leads to the concept of a topological group, and the second level is a differentiable structure, which leads to the concept of a Lie group.

Definition 4.1.8. (a) A *Lie group* G is a smooth manifold endowed with a group structure such that the multiplication and inversion map

$$m_G \colon G \times G \to G \quad \text{and} \quad \eta_G \colon G \to G$$

are smooth. Since smooth maps are continuous, every Lie group is in particular a topological group.

(b) If G and H are topological (Lie) groups, then a group homomorphism $\varphi \colon G \to H$ is called a *morphism of topological (Lie) groups* if φ is continuous (smooth).

Remark 4.1.9. As for topological groups, it is easy to see that the smoothness requirements in the definition of a Lie group are equivalent to the requirement that the map

$$G \times G \to G, \quad (x, y) \mapsto xy^{-1}$$

is smooth.

Lemma 4.1.10. *Let G be a Lie group and $g \in G$. Then the following maps are diffeomorphisms of G:*

(1) $\lambda_g \colon G \to G, x \mapsto gx$ *(left translations).*

(2) $\rho_g \colon G \to G, x \mapsto xg$ *(right translations).*

(3) $c_g \colon G \to G, x \mapsto gxg^{-1}$ *(conjugations).*

Proof. The smoothness of all these maps follows from the smoothness of the group operations. That they are diffeomorphisms is a consequence of their bijectivity and $\lambda_g^{-1} = \lambda_{g^{-1}}$, $\rho_g^{-1} = \rho_{g^{-1}}$ and $c_g^{-1} = c_{g^{-1}}$. □

Example 4.1.11. (Vector groups) Each Banach space E is an abelian Lie group with respect to addition and the obvious manifold structure (Example 4.1.3(a)).

Vector groups $(E, +)$ form the most elementary Lie groups. The next natural class are unit groups of Banach algebras.

Example 4.1.12. (Unit groups as Lie groups) Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{K}$ and $\mathcal{A}^\times$ be its unit group. As an open subset of $\mathcal{A}$, the group $\mathcal{A}^\times$ carries a natural manifold structure. The multiplication on $\mathcal{A}$ is bilinear and continuous, hence a smooth map (cf. Lemma B.2.3). Therefore the multiplication of $\mathcal{A}^\times$ is smooth and it remains to see that the inversion $\eta \colon \mathcal{A}^\times \to \mathcal{A}^\times$ is smooth. We know already from Proposition 1.1.9 that η is continuous.

For $a, b \in \mathcal{A}^\times$, we have $b^{-1} - a^{-1} = a^{-1}(a - b)b^{-1}$, which implies that

$$\eta(a + h) - \eta(a) = (a + h)^{-1} - a^{-1} = a^{-1}(-h)(a + h)^{-1} = -a^{-1}h(a + h)^{-1}.$$

This implies that the directional derivative

$$\mathtt{d}\eta(a)(h) = \frac{d}{dt}\Big|_{t=0} \eta(a+th) = \lim_{t\to 0} -a^{-1}h(a+th)^{-1} = -a^{-1}ha^{-1} \tag{4.2}$$

exists. Moreover,

$$\begin{aligned} &\frac{1}{\|h\|}\big\|\eta(a+h) - \eta(a) + a^{-1}ha^{-1}\big\| \\ &= \frac{1}{\|h\|}\big\|a^{-1}h\big((a+h)^{-1} - a^{-1}\big)\big\| \le \|a^{-1}\|\|(a+h)^{-1} - a^{-1}\| \to 0 \end{aligned}$$

for $h \to 0$ follows from the continuity of η on $\mathcal{A}^\times$. Therefore η is differentiable in a with

$$\mathtt{d}\eta(a)(h) = -a^{-1}ha^{-1}. \tag{4.3}$$

The continuity of η implies that $\mathtt{d}\eta(a) = -\lambda_{a^{-1}}\rho_{a^{-1}} = -\lambda_{\eta(a)}\rho_{\eta(a)} : \mathcal{A}^\times \to \mathcal{L}(\mathcal{A})$ is continuous, hence that η is a C^1-map. From (4.3) and the smoothness of the continuous bilinear map

$$\mathcal{A}^2 \to \mathcal{L}(\mathcal{A}), \quad (x,y) \mapsto \lambda_x\rho_y,$$

we further derive that, if η is C^k, then $\mathtt{d}\eta$ is also C^k, so that η is C^{k+1}. Inductively, it follows that η is smooth.

Exercises for Section 4.1

Exercise 4.1.1. Let G and H be Lie groups and $\varphi: G \to H$ be a group homomorphism. Show that φ is smooth if there exists an open identity neighborhood $U \subseteq G$ on which φ is smooth.

Exercise 4.1.2. Let G and H be Lie groups and $\varphi: G \to H$ be a bijective group homomorphism. Show that φ is a diffeomorphism if there exists an open $\mathbf{1}$-neighborhood $U \subseteq G$ such that $\varphi|_U$ is a diffeomorphism onto an open $\mathbf{1}$-neighborhood in H.

4.2 Constructing Lie Group Structures on Groups

In this subsection we describe some methods to construct Lie group structures on groups, starting from a manifold structure on some "identity neighborhood" for which the group operations are smooth close to $\mathbf{1}$.

Lie groups from local data

The following theorem is the smooth version of Lemma 1.4.5 from the topological context. It is our main tool to construct Lie group structures on groups.

Theorem 4.2.1. *Let G be a group and $U = U^{-1}$ a symmetric subset containing $\mathbf{1}$. We further assume that U is a smooth manifold and that*

(L1) *$D := \{(x,y) \in U \times U : xy \in U\}$ is an open subset and the multiplication $m_U : D \to U, (x,y) \mapsto xy$ is smooth,*

(L2) *the inversion map $\eta_U : U \to U, u \mapsto u^{-1}$ is smooth, and,*

(L3) *for each $g \in G$, there exists an open $\mathbf{1}$-neighborhood $U_g \subseteq U$ with $c_g(U_g) \subseteq U$ and such that the conjugation map*

$$c_g : U_g \to U, \quad x \mapsto gxg^{-1}$$

is smooth.

Then there exists a unique structure of a Lie group on G such that the inclusion map $U \hookrightarrow G$ is a diffeomorphism onto an open subset of G.

If, in addition, U generates G, then (L1/2) *imply* (L3).

Proof. From Lemma 1.4.5, we obtain a unique group topology on G for which the inclusion map $U \hookrightarrow G$ is an open embedding.

Now we turn to the manifold structure. Let $V = V^{-1} \subseteq U$ be an open $\mathbf{1}$-neighborhood with $VV \times VV \subseteq D$, for which there exists a Banach space E and an E-chart (φ, V) of U. For $g \in G$ we consider the map

$$\varphi_g : gV \to E, \quad \varphi_g(x) = \varphi(g^{-1}x)$$

which is a homeomorphisms of the open subset gV of G onto the open subset $\varphi(V) \subseteq E$. We claim that $(\varphi_g, gV)_{g\in G}$ is a smooth atlas of G.

Let $g_1, g_2 \in G$ and put $W := g_1V \cap g_2V$. If $W \neq \emptyset$, then $g_2^{-1}g_1 \in VV^{-1} = VV$. The smoothness of the map

$$\psi := \varphi_{g_2} \circ \varphi_{g_1}^{-1}|_{\varphi_{g_1}(W)} : \varphi_{g_1}(W) \to \varphi_{g_2}(W)$$

given by

$$\psi(x) = \varphi_{g_2}(\varphi_{g_1}^{-1}(x)) = \varphi_{g_2}(g_1\varphi^{-1}(x)) = \varphi(g_2^{-1}g_1\varphi^{-1}(x))$$

follows from the smoothness of the multiplication $VV \times VV \to U$. This proves that the charts $(\varphi_g, gU)_{g\in G}$ form a smooth atlas of G. Moreover, the construction implies that all left translations of G are smooth maps because $\varphi_g \circ \lambda_h = \varphi_{g^{-1}h}$ for $g, h \in G$.

The construction also shows that, for $g \in G$, the conjugation $c_g : G \to G$ is smooth in a neighborhood of $\mathbf{1}$. Since all left translations are smooth, and

$$c_g \circ \lambda_x = \lambda_{c_g(x)} \circ c_g$$

is smooth in an identity neighborhood, the smoothness of c_g in a neighborhood of $x \in G$ follows. Therefore all conjugations and hence also all right multiplications are smooth. The smoothness of the inversion follows from its smoothness on V and the fact that left and right multiplications are smooth because

$$\eta_G \circ \lambda_g = \rho_g^{-1} \circ \eta_G$$

is smooth in a neighborhood of $\mathbf{1}$. Finally, the smoothness of the multiplication follows from the smoothness in a neighborhood of $(\mathbf{1}, \mathbf{1})$ because it implies that

$$m_G = \lambda_g \circ \rho_h \circ m_G \circ (\lambda_{g^{-1}} \times \rho_{h^{-1}})$$

is smooth in a neighborhood of (g, h). We conclude that G is a Lie group.

Next we show that the inclusion $U \hookrightarrow G$ of U is a diffeomorphism. So let $x \in U$ and recall the open set $U_x = U \cap x^{-1}U$. Then λ_x restricts to a smooth map $U_x \to U$ with image $U_{x^{-1}}$. Its inverse is also smooth. Hence $\lambda_x^U \colon U_x \to U_{x^{-1}}$ is a diffeomorphism. Since $\lambda_x \colon G \to G$ also is a diffeomorphism, it follows that the inclusion $\lambda_x \circ \lambda_{x^{-1}}^U \colon U_{x^{-1}} \to G$ is a diffeomorphism. As x was arbitrary, the inclusion of U in G is a diffeomorphic embedding.

The uniqueness of the Lie group structure is clear because each locally diffeomorphic bijective homomorphism between Lie groups is a diffeomorphism (cf. Exercise 4.1.2).

Finally, we assume that G is generated by U. We show that in this case (L3) is a consequence of (L1) and (L2). For each $g \in U$, there exists an open $\mathbf{1}$-neighborhood U_g with $gU_g \times \{g^{-1}\} \subseteq D$. Then $c_g(U_g) \subseteq U$, and the smoothness of m_U implies that $c_g|_{U_g} \colon U_g \to U$ is smooth. Hence, for each $g \in U$, the conjugation c_g is smooth in a neighborhood of $\mathbf{1}$. Since the set of all these g is a submonoid of G containing U, it contains U^n for each $n \in \mathbb{N}$, hence all of G because G is generated by $U = U^{-1}$. Therefore (L3) is satisfied. □

Remark 4.2.2. Suppose that G is a Lie group. Let (U, φ) be a chart with $\mathbf{1} \in U = U^{-1}$. Then $\varphi \colon U \to \varphi(U)$ is a homeomorphism onto an open subset of some Banach space E. Let $V \subseteq U$ be a symmetric $\mathbf{1}$-neighborhood with $VV \subseteq U$. Then the map

$$\varphi(V) \times \varphi(V) \to \varphi(U), \quad (x, y) \mapsto \varphi(\varphi^{-1}(x)\varphi^{-1}(y))$$

is smooth. This is the kind of situation one has in mind in the preceding theorem.

The main point in the local approach is that it emphasizes that the whole differentiable structure is determined by the manifold structure on a suitable neighborhood of the identity, which is very convenient to construct Lie group structures.

Corollary 4.2.3. *Let G be a group and $N \trianglelefteq G$ be a normal subgroup of G that carries a Lie group structure. Then there exists a Lie group structure on G for which N is an open subgroup if and only if, for each $g \in G$, the restriction $c_g|_N$ is a smooth automorphism of N.*

Proof. If N is an open normal subgroup of the Lie group G, then clearly all inner automorphisms of G restrict to smooth automorphisms of N.

Suppose, conversely, that N is a normal subgroup of the group G which is a Lie group and that all inner automorphisms of G restrict to smooth automorphisms of N. Then we can apply Theorem 4.2.1 with $U = N$ and obtain a Lie group structure on G for which the inclusion $N \to G$ is a diffeomorphism onto an open subgroup of G. □

Linear groups as Lie groups

Before we turn to linear Lie groups, we explain how any closed Lie subalgebra $\mathfrak{g}$ can be used to construct a Lie group modeled on $\mathfrak{g}$.

Theorem 4.2.4. (Integral Subgroup Theorem; linear version) *Let $\mathcal{A}$ be a unital Banach algebra and $\mathfrak{g} \subseteq \mathcal{A}$ be a closed Lie subalgebra. Then the subgroup $G := \langle \exp \mathfrak{g} \rangle$ of $\mathcal{A}^\times$ generated by $\exp(\mathfrak{g})$ carries a Lie group structure for which there exists an open 0-neighborhood $V \subseteq \mathfrak{g}$ on which the Dynkin series converges and*

$$\exp\colon \mathfrak{g} \to G, \quad x \mapsto \exp x$$

is smooth and maps V diffeomorphism onto its open image in G and satisfies

$$\exp(x * y) = \exp(x)\exp(y) \quad \textit{for} \quad x, y \in V.$$

Proof. Let

$$V := \Big\{x \in \mathfrak{g}\colon \|x\| < \tfrac{1}{2}\log\Big(2 - \frac{\sqrt{2}}{2}\Big)\Big\},$$

so that the Dynkin series for $x * y$ converges for $x, y \in V$ and satisfies

$$\exp(x * y) = \exp(x)\exp(y)$$

(Proposition 2.3.6). Since $x * y$ is a series whose summands are obtained by iterated Lie brackets, the closedness of $\mathfrak{g}$ implies that $x * y \in \mathfrak{g}$ for $x, y \in V$. The function $V \times V \to \mathfrak{g}, (x, y) \mapsto x * y$ is smooth because it is the restriction of a smooth function on an open 0-neighborhood of $\mathcal{A}$.

We consider the subset $U := \exp(V) \subseteq G$. From $V = -V$ we derive $U = U^{-1}$, and since $\|x\| < \log 2$ for $x \in V$, Proposition 2.2.3 implies that $\varphi := \exp|_V$ is injective. We may thus endow U with the manifold structure turning φ into a diffeomorphism.

Then

$$D_V := \{(x, y) \in V \times V \colon x * y \in V\}$$

is an open subset of $V \times V$ on which the BCDH multiplication is smooth, so that the multiplication

$$D_U := \{(g, h) \in U \times U \colon gh \in U\} = \{(\exp x, \exp y)\colon (x, y) \in D_V\} \to U$$

is also smooth. We further observe that

$$\exp(-x) = \exp(x)^{-1},$$

from which it follows that the inversion on U is smooth.

To verify (L3), we first note that, for each $x \in \mathfrak{g}$, we have $\mathrm{Ad}(\exp x)\mathfrak{g} = e^{\mathrm{ad}\, x}\mathfrak{g} \subseteq \mathfrak{g}$, from which it follows that $\mathrm{Ad}(g)\mathfrak{g} = \mathfrak{g}$ for each $g \in G$. Hence $\mathrm{Ad}(g)$ induces a topological linear automorphism of $\mathfrak{g}$, so that there exists an open 0-neighborhood $V_g \subseteq \mathfrak{g}$ with $\mathrm{Ad}(g)V_g \subseteq V$. Now $\mathrm{Ad}(g)$ restrict to a smooth map

$$\mathrm{Ad}(g)\colon V_g \to V, \quad x \mapsto gxg^{-1}$$

with $\varphi(\mathrm{Ad}(g)x) = g\varphi(x)g^{-1}$.

Therefore U satisfies all assumptions of Theorem 4.2.1, so that we obtain a Lie group structure on G for which φ, resp., exp induces a local diffeomorphism in 0.

Finally, the smoothness of the exponential function follows from its smoothness on V, $\bigcup_{m\in\mathbb{N}} mV = \mathfrak{g}$, and $\exp_G(mx) = \exp_G(x)^m$. $\square$

Remark 4.2.5. The example of the dense wind shows that we cannot expect that the group $G = \langle \exp \mathfrak{g}\rangle$ is closed in $\mathcal{A}^\times$ or that the inclusion map $G \to \mathcal{A}^\times$ is a topological embedding. However, the smoothness of the exponential map $\mathfrak{g} \to \mathcal{A}^\times$ and the fact that $\exp|_{\mathfrak{g}} \colon \mathfrak{g} \to G$ is a local diffeomorphism imply that the inclusion $G \to \mathcal{A}^\times$ is smooth, i.e., a morphism of Lie groups (Exercise 4.1.2).

Theorem 4.2.6. (Linear Lie Group Theorem) *Every linear Lie group carries a unique Lie group structure for which the exponential function* $\exp_G \colon \mathbf{L}(G) \to G$ *is smooth and there exists an open* 0*-neighborhood* $V \subseteq \mathbf{L}(G)$ *on which the Dynkin series converges and* $\exp_G|_V$ *is a diffeomorphism onto an open subset of* G *with*

$$\exp_G(x * y) = \exp_G(x)\exp_G(y) \quad \textit{for} \quad x, y \in V.$$

Proof. Let $\mathcal{A}$ be a unital Banach algebra and $G \subseteq \mathcal{A}^\times$ be a linear Lie group. By definition, there exists an open 0-neighborhood $W \subseteq \mathbf{L}(G)$ for which $\varphi := \exp_G|_W$ is a homeomorphism onto an open $\mathbf{1}$-neighborhood in G. In particular, $\exp(\mathbf{L}(G)) \subseteq G$ is a connected 0-neighborhood, so that $\langle \exp \mathbf{L}(G)\rangle$ is an open connected subgroup of G, hence coincides with the identity component G_0 of G (cf. Lemmas 1.2.2(iv) and 1.2.10). We endow G_0 with the Lie group structure from Theorem 4.2.4. Since exp maps a 0-neighborhood of $\mathbf{L}(G)$ homeomorphically onto a $\mathbf{1}$-neighborhood of G_0 and likewise onto a $\mathbf{1}$-neighborhood of G, it follows that the Lie group structure on G_0 is compatible with the group topology on G because both have the same $\mathbf{1}$-neighborhoods.

To see that the Lie group structure on G_0 extends to G, we have to verify that all conjugation maps c_g, $g \in G$, restrict to smooth maps on G_0 (Corllary 4.2.3). This follows from

$$c_g(\exp x) = \exp(\mathrm{Ad}(g)x) \quad \text{for} \quad x \in \mathbf{L}(G)$$

and Exercise 4.1.1 because the linear maps $\mathrm{Ad}(g) \colon \mathbf{L}(G) \to \mathbf{L}(G)$ are continuous, hence smooth. The uniqueness of this Lie group structure now follows from Exercise 4.1.2. $\square$

Remark 4.2.7. If $G \subseteq \mathcal{A}^\times$ is a closed subgroup, not necessarily Lie, then its Lie algebra $\mathbf{L}(G)$ is a closed Lie subalgebra of $\mathcal{A}_L$ and Theorem 4.2.4 yields a Lie group structure on the subgroup $G_1 := \langle \exp \mathbf{L}(G)\rangle$ of G which the exponential map $\exp \colon L(G) \to G_1$ is a local homeomorphism. The relation $c_g(\exp x) = \exp(\mathrm{Ad}(g)x)$ implies that $G_1 \trianglelefteq G$ is a normal subgroup and that the conjugation maps $c_g \colon G_1 \to G_1$, $g \in G$, are smooth automorphisms of G_1. Therefore Corollary 4.2.3 further implies that G carries a unique Lie group structure for which G_1 is an open subgroup. The so obtained Lie group structure is compatible with the subspace topology on G inherited from $\mathcal{A}^\times$ if and only if G is a linear Lie group.

4.3 Coverings of Lie groups

In the preceding section we have seen how to construct Lie group structures on groups from local data. This construction applies in particular to those quotient morphisms $q\colon G \to G/N$, where G is a Lie group and q is a local homeomorphism, i.e., maps some open identity neighborhood homeomorphically to an open identity neighborhood in N. This means that N is a discrete subgroup of G, such as $\mathbb{Z}$ in $\mathbb{R}$ (cf. Exercise 4.3.2).

To deal properly with such maps, we recall the concept of a covering map from Definition C.2.1. This concept is particularly important in the theory of Lie groups because it can be used to understand how different connected Lie groups with the same Lie algebra can be.

The following lemma shows that covering morphism of topological groups coincide with quotient maps modulo discrete normal subgroups.

Lemma 4.3.1. (a) *If $\varphi\colon G \to H$ is a covering morphism of topological groups, then its kernel $\Gamma := \ker\varphi$ is a discrete normal subgroup and the induced homomorphism $\overline{\varphi}\colon G/\Gamma \to H$ is an isomorphism of topological groups.*
(b) *Conversely, for every discrete normal subgroup Γ of a topological group G, the quotient map $q\colon G \to G/\Gamma$ is a covering morphism.*

Proof. (a) Since φ is a covering, there exists an open $\mathbf{1}$-neighborhood $U \subseteq G$ which is mapped homeomorphically onto $\varphi(U)$. Then

$$\Gamma \cap U = \{u \in U : \varphi(u) = \mathbf{1}\} = \{\mathbf{1}\}$$

implies that Γ is discrete (cf. Remark 1.2.3).

The induced map $\overline{\varphi}\colon G/\Gamma \to H$ is a bijective continuous homomorphism of topological groups, and since φ is an open map, the same holds for $\overline{\varphi}$ (cf. Lemma 1.4.7). Therefore $\overline{\varphi}$ is a homeomorphism.

(b) Let $U \subseteq G$ be an open symmetric $\mathbf{1}$-neighborhood with $U^2 \cap \Gamma = \{1\}$. Then, for each $g \in G$, $gU\Gamma = \bigcup_{\gamma\in\Gamma} gU\gamma$ is a disjoint union and an open subset of G. Since $gu_1\Gamma = gu_2\Gamma$ for $u_j \in U$ implies that $u_2^{-1}u_1 \in U^2 \cap \Gamma = \{\mathbf{1}\}$, the restriction of q to each set $gU\gamma$ is injective, hence a homeomorphism onto the open subset $q(gU)$ of G/Γ. In view of $gU\Gamma = q^{-1}(q(gU))$, it follows that q is a covering. □

Proposition 4.3.2. *Let $\varphi\colon G \to H$ be a covering of topological groups. If G or H is a Lie group, then the other group has a unique Lie group structure for which φ is a morphism of Lie groups which is a local diffeomorphism.*

Proof. Let $U_G \subseteq G$ be an open symmetric $\mathbf{1}$-neighborhood for which $\varphi|_{U_G}$ is a homeomorphism onto an open subset U_H of H and which satisfies $U_G^3 \cap \ker\varphi = \{\mathbf{1}\}$.

Suppose first that G is a Lie group. Then we apply Theorem 4.2.1 to U_H, endowed with the manifold structure for which $\varphi|_{U_G}$ is a diffeomorphism. Then (L2) follows from $\varphi(x)^{-1} = \varphi(x^{-1})$. To verify the smoothness of the multiplication map

$$m_{U_H}\colon D_H := \{(a,b) \in U_H \times U_H : ab \in U_H\} \to U_H,$$

we first observe that, if $x, y \in U_G$ satisfy $(\varphi(x), \varphi(y)) \in D_H$, i.e., $\varphi(xy) \in U_H$, then there exists a $z \in U_G$ with $\varphi(xy) = \varphi(z)$, and $xyz^{-1} \in U_G{}^3 \cap \ker(\varphi) = \{\mathbf{1}\}$ yields $xy = z \in U_G$. We thus have $D_H = (\varphi \times \varphi)(D_G)$ for

$$D_G := \{(x, y) \in U_G \times U_G \colon xy \in U_G\}$$

and the smoothness of m_H follows from the smoothness of the multiplication $m_{U_G} \colon D_G \to U_G$ and

$$m_{U_H} \circ (\varphi \times \varphi) = \varphi \circ m_{U_G}.$$

To verify (L3), we note that the surjectivity of φ implies that for each $h \in H$ there is an element $g \in G$ with $\varphi(g) = h$. Now we choose an open $\mathbf{1}$-neighborhood $U_g \subseteq U_G$ with $c_g(U_g) \subseteq U_G$ and put $U_h := \varphi(U_g)$.

If, conversely, H is a Lie group, then we apply Theorem 4.2.1 to U_G, endowed with the manifold structure for which $\varphi|_{U_G}$ is a diffeomorphism onto U_H. Again, (L2) follows right away, and (L1) follows from $(\varphi \times \varphi)(D_G) \subseteq D_H$ and the smoothness of

$$m_{U_H} \circ (\varphi \times \varphi) = \varphi \circ m_{U_G}.$$

For (L3), we choose U_g as any open $\mathbf{1}$-neighborhood in U_G with $c_g(U_g) \subseteq U$. Then the smoothness of $c_g|_{U_g}$ follows from the smoothness the maps of $\varphi \circ c_g = c_{\varphi(g)} \circ \varphi$. □

Theorem 4.3.3. (Universal Covering Theorem) *If G is a connected Lie group, then there exists a covering map $q_G \colon \widetilde{G} \to G$, where $\widetilde{G}$ is a simply connected Lie group.*

Proof. Since every $g \in G$ has a neighborhood U homeomorphic to a ball in a Banach space and balls are simply connected, G is locally arcwise connected and semilocally simply connected. Therefore Theorem C.2.13 implies the existence of a covering $q_G \colon \widetilde{G} \to G$ by a simply connected space $\widetilde{G}$.

Next we construct a (topological) group structure on $\widetilde{G}$. Pick an element $\widetilde{\mathbf{1}} \in q_G^{-1}(\mathbf{1})$. Then the multiplication map $m_G \colon G \times G \to G$ resp. the map $m_G \circ (q_G \times q_G)$ lifts uniquely to a continuous map $\widetilde{m}_G \colon \widetilde{G} \times \widetilde{G} \to \widetilde{G}$ with $\widetilde{m}_G(\widetilde{\mathbf{1}}, \widetilde{\mathbf{1}}) = \widetilde{\mathbf{1}}$ (The Lifting Theorem C.2.9). To see that the multiplication map $\widetilde{m}_G$ is associative, we observe that

$$\begin{aligned} & q_G \circ \widetilde{m}_G \circ (\mathrm{id}_{\widetilde{G}} \times \widetilde{m}_G) = m_G \circ (q_G \times q_G) \circ (\mathrm{id}_{\widetilde{G}} \times \widetilde{m}_G) \\ = \;& m_G \circ (\mathrm{id}_G \times m_G) \circ (q_G \times q_G \times q_G) = m_G \circ (m_G \times \mathrm{id}_G) \circ (q_G \times q_G \times q_G) \\ = \;& q_G \circ \widetilde{m}_G \circ (\widetilde{m}_G \times \mathrm{id}_{\widetilde{G}}), \end{aligned}$$

so that the two continuous maps

$$\widetilde{m}_G \circ (\mathrm{id}_{\widetilde{G}} \times \widetilde{m}_G), \quad \widetilde{m}_G \circ (\widetilde{m}_G \times \mathrm{id}_{\widetilde{G}}) \colon \widetilde{G}^3 \to G,$$

are lifts of the same map $G^3 \to G$ and both map $(\widetilde{\mathbf{1}}, \widetilde{\mathbf{1}}, \widetilde{\mathbf{1}})$ to $\widetilde{\mathbf{1}}$. Hence the uniqueness of lifts implies that $\widetilde{m}_G$ is associative. We likewise obtain that the unique lift $\widetilde{\eta}_G \colon \widetilde{G} \to \widetilde{G}$ of the inversion map $\eta_G \colon G \to G$ with $\widetilde{\eta}_G(\widetilde{\mathbf{1}}) = \widetilde{\mathbf{1}}$ satisfies

$$\widetilde{m}_G \circ (\eta_G \times \mathrm{id}_{\widetilde{G}}) = \widetilde{\mathbf{1}} = \widetilde{m}_G \circ (\mathrm{id}_{\widetilde{G}} \times \eta_G).$$

Therefore $\widetilde{m}_G$ defines on $\widetilde{G}$ a topological group structure such that $q_G\colon \widetilde{G} \to G$ is a covering morphism of topological groups. Now Proposition 4.3.2 applies. □

Theorem 4.3.4. (Lifting Theorem for Groups) *Let $q\colon G \to H$ be a covering morphism of Lie groups. If $f\colon L \to H$ is a morphism of Lie groups, where L is simply connected, then there exists a unique morphism of Lie groups $\widetilde{f}\colon L \to G$ with $q \circ \widetilde{f} = f$.*

Proof. Since Lie groups are locally arcwise connected, the Lifting Theorem C.2.9 implies the existence of a unique lift $\widetilde{f}\colon L \to G$ with $\widetilde{f}(\mathbf{1}_L) = \mathbf{1}_G$. Then

$$m_G \circ (\widetilde{f} \times \widetilde{f})\colon L \times L \to G$$

is the unique lift of $m_H \circ (f \times f)\colon L \times L \to H$ mapping $(\mathbf{1}_L, \mathbf{1}_L)$ to $\mathbf{1}_G$. We also have

$$q \circ \widetilde{f} \circ m_L = f \circ m_L = m_H \circ (f \times f),$$

so that $\widetilde{f} \circ m_L$ is another lift of $m_H \circ (f \times f)$, mapping $(\mathbf{1}_L, \mathbf{1}_L)$ to $\mathbf{1}_G$. Therefore

$$\widetilde{f} \circ m_L = m_G \circ (\widetilde{f} \times \widetilde{f}),$$

which means that $\widetilde{f}$ is a group homomorphism.

Since q is a local diffeomorphism and $\widetilde{f}$ is a continuous lift of f, it is also smooth in an identity neighborhood of L, hence smooth (Exercise 4.1.1; see also Exercise 4.3.3). □

Theorem 4.3.5. *Let G be a connected Lie group and $q_G\colon \widetilde{G} \to G$ be a universal covering homomorphism. Then $\ker q_G \cong \pi_1(G)$ is a discrete central subgroup and $G \cong \widetilde{G}/\ker q_G$.*

Moreover, for any discrete central subgroup $\Gamma \subseteq \widetilde{G}$, the group $\widetilde{G}/\Gamma$ is a connected Lie group with the same universal covering group as G. We thus obtain a bijection from the set of all $\mathrm{Aut}(\widetilde{G})$-orbits in the set of discrete central subgroups of $\widetilde{G}$ onto the set of isomorphy classes of connected Lie groups whose universal covering group is isomorphic to $\widetilde{G}$.

Proof. First we note that $\ker q_G$ is a discrete normal subgroup of the connected Lie group $\widetilde{G}$, hence central by Exercise 1.2.8. Left multiplications by elements of $\ker q_G$ lead to deck transformations of the covering $\widetilde{G} \to G$, and this group of deck transformations acts transitively on the fiber $\ker q_G$ of $\mathbf{1}$. Proposition C.2.17 now yields a group isomorphism

$$\pi_1(G) \cong \ker q_G \tag{4.4}$$

Since $q_G\colon \widetilde{G} \to G$ is a covering which is a local diffeomorphism, the induced map $\widetilde{G}/\ker q_G \to G$ also is a local diffeomorphism if $\widetilde{G}/\ker q_G$ carries the canonical Lie group structure (Proposition 4.3.2). Since it also is bijective, it is an isomorphism of Lie groups (Exercise 4.1.2).

If, conversely, $\Gamma \subseteq \widetilde{G}$ is a discrete central subgroup, then $\widetilde{G}/\Gamma$ is a Lie group whose universal covering group is $\widetilde{G}$ (Proposition 4.3.2). Two such groups $\widetilde{G}/\Gamma_1$ and $\widetilde{G}/\Gamma_2$

are isomorphic if and only if there exists a Lie group automorphism $\varphi \in \mathrm{Aut}(\widetilde{G})$ with $\varphi(\Gamma_1) = \Gamma_2$ (Theorem 4.3.4). Therefore the isomorphism classes of Lie groups with the same universal covering group G are parameterized by the orbits of the group $\mathrm{Aut}(\widetilde{G})$ in the set of discrete central subgroups of $\widetilde{G}$. □

Examples 4.3.6. Let E be a Banach space, so that $(E, +)$ is a Lie group. Then, for each discrete subgroup $\Gamma \subseteq E$, the quotient group E/Γ carries a natural Lie group structure (Proposition 4.3.2; Exercise 4.3.2). Since E is simply connected by Remark C.1.7, E is the universal covering group of E/Γ and $\pi_1(E/\Gamma) \cong \Gamma$ (Theorem 4.3.5). We shall see in Corollary 6.2.19 below that all connected abelian Lie groups are of this form. As special cases we obtain in particular the finite-dimensional tori

$$\mathbb{T}^d \cong \mathbb{R}^d/\mathbb{Z}^d$$

(cf. Example 1.4.8).

To classify all connected abelian Lie groups A with $\widetilde{A} \cong \mathbb{R}^n$, we can now proceed as follows. Since all automorphisms of the topological group $\mathbb{R}^n$ are automatically linear (Exercise 4.3.5), we have an isomorphism $\mathrm{Aut}(\widetilde{A}) \cong \mathrm{GL}_n(\mathbb{R})$. Further, Exercise 1.2.7 implies that the discrete subgroup $\ker q_A \cong \pi_1(A)$ of $\widetilde{A} \cong \mathbb{R}^n$ can be mapped by some $\varphi \in \mathrm{GL}_n(\mathbb{R})$ onto

$$\mathbb{Z}^k \cong \mathbb{Z}^k \times \{0\} \subseteq \mathbb{R}^k \times \mathbb{R}^{n-k} \cong \mathbb{R}^n.$$

Therefore

$$A \cong \mathbb{R}^n/\mathbb{Z}^k \cong \mathbb{T}^k \times \mathbb{R}^{n-k},$$

and it is clear that the number k is an isomorphy invariant of the Lie group A, namely, the rank of its fundamental subgroup.

Since we shall see below that every connected abelian Lie groups A of dimension n satisfies $\widetilde{A} \cong \mathbb{R}^n$, it follows that A is determined up to isomorphism by the pair (n, k), where $n = \dim A$ and $k = \operatorname{rank} \pi_1(A)$. The case $n = k$ leads to the compact connected abelian Lie groups.

Examples 4.3.7. (a) The group $\mathbb{T} \cong \mathbb{R}/\mathbb{Z}$ is homeomorphic to the one-dimensional sphere $\mathbb{S}^1$, which is not simply connected.

The group

$$\mathrm{SU}_2(\mathbb{C}) \cong \left\{ \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \in \mathrm{GL}_2(\mathbb{C})\colon |a|^2 + |b|^2 = 1 \right\}$$

is homeomorphic to the 3-sphere

$$\{(a, b) \in \mathbb{C}^2 \colon \|(a, b)\| = 1\} \cong \mathbb{S}^3$$

which is simply connected (Exercise C.1.3). One can show that the sphere $\mathbb{S}^n$ carries a Lie group structure if and only if $n = 0, 1, 3$.

(b) With some more advanced tools from homotopy theory, one can show that the groups $\mathrm{SU}_n(\mathbb{C})$ are always simply connected. However, this is never the case for the groups $\mathrm{U}_n(\mathbb{C})$.

To see this, consider the group homomorphism

$$\gamma\colon \mathbb{T} \to \mathrm{U}_n(\mathbb{C}), \quad z \mapsto \mathrm{diag}(z, 1, \ldots, 1)$$

and note that $\det \circ \gamma = \mathrm{id}_{\mathbb{T}}$. From that one easily derives that the multiplication map

$$\mu\colon \mathrm{SU}_n(\mathbb{C}) \times \mathbb{T} \to \mathrm{U}_n(\mathbb{C}), \quad (g, z) \mapsto g\gamma(z)$$

is a homeomorphism, so that

$$\pi_1(\mathrm{U}_n(\mathbb{C})) \cong \pi(\mathrm{SU}_n(\mathbb{C})) \times \pi_1(\mathbb{T}) \cong \pi_1(\mathbb{T}) \cong \mathbb{Z}.$$

We further derive that the universal covering group is given by

$$\widetilde{\mathrm{U}}_n(\mathbb{C}) \cong \mathrm{SU}_n(\mathbb{C}) \rtimes_\beta \mathbb{R} \quad \text{where} \quad \beta(t)g := \gamma(e^{it})g\gamma(e^{-it}).$$

Example 4.3.8. We show that

$$\pi_1(\mathrm{SO}_3(\mathbb{R})) \cong C_2 = \{\pm 1\}$$

by constructing a surjective homomorphism

$$\varphi\colon \mathrm{SU}_2(\mathbb{C}) \to \mathrm{SO}_3(\mathbb{R})$$

with $\ker \varphi = \{\pm \mathbf{1}\}$, so that

$$\mathrm{SO}_3(\mathbb{R}) \cong \mathrm{SU}_2(\mathbb{C})/\{\pm \mathbf{1}\}.$$

Since $\mathrm{SU}_2(\mathbb{C})$ is homeomorphic to $\mathbb{S}^3$, it is simply connected (Exercise C.1.3), so that we obtain $\pi_1(\mathrm{SO}_3(\mathbb{R})) \cong C_2$ (Theorem 4.3.5).

We consider

$$\mathfrak{su}_2(\mathbb{C}) = \{x \in \mathfrak{gl}_2(\mathbb{C})\colon x^* = -x, \mathrm{tr}\, x = 0\} = \left\{ \begin{pmatrix} ai & b \\ -\overline{b} & -ai \end{pmatrix} : b \in \mathbb{C}, a \in \mathbb{R} \right\}$$

and observe that this is a three-dimensional real subspace of $\mathfrak{gl}_2(\mathbb{C})$. We obtain on $E := \mathfrak{su}_2(\mathbb{C})$ the structure of a euclidean vector space by the scalar product

$$\beta(x, y) := \mathrm{tr}(xy^*) = -\mathrm{tr}(xy).$$

Now we consider the adjoint representation

$$\mathrm{Ad}\colon \mathrm{SU}_2(\mathbb{C}) \to \mathrm{GL}(E), \quad \mathrm{Ad}(g)(x) = gxg^{-1}.$$

Then we have for $x, y \in E$ and $g \in \mathrm{SU}_2(\mathbb{C})$ the relation

$$\begin{aligned} \beta\big(\mathrm{Ad}(g)x, \mathrm{Ad}(g)y\big) &= \mathrm{tr}(gxg^{-1}(gyg^{-1})^*) = \mathrm{tr}(gxg^{-1}(g^{-1})^*y^*g^*) \\ &= \mathrm{tr}(gxg^{-1}gy^*g^{-1}) = \mathrm{tr}(xy^*) = \beta(x, y). \end{aligned}$$

This means that

$$\mathrm{Ad}(\mathrm{SU}_2(\mathbb{C})) \subseteq \mathrm{O}(E, \beta) \cong \mathrm{O}_3(\mathbb{R}).$$

Since $\mathrm{SU}_2(\mathbb{C})$ is connected, we further obtain $\mathrm{Ad}(\mathrm{SU}_2(\mathbb{C})) \subseteq \mathrm{SO}(E,\beta) \cong \mathrm{SO}_3(\mathbb{R})$, the identity component of $\mathrm{O}(E,\beta)$.

The derived representation is given by

$$\mathbf{L}(\mathrm{Ad}) = \mathrm{ad}\colon \mathfrak{su}_2(\mathbb{C}) \to \mathfrak{so}(E,\beta) \cong \mathfrak{so}_3(\mathbb{R}), \quad \mathrm{ad}(x)(y) = [x,y].$$

If $\mathrm{ad}\, x = 0$, then $\mathrm{ad}\, x(i\mathbf{1}) = 0$ implies that $\mathrm{ad}\, x(\mathfrak{u}_2(\mathbb{C})) = \{0\}$, so that $\mathrm{ad}\, x(\mathfrak{gl}_2(\mathbb{C})) = \{0\}$ follows from $\mathfrak{gl}_2(\mathbb{C}) = \mathfrak{u}_2(\mathbb{C}) + i\mathfrak{u}_2(\mathbb{C})$. This implies that $x \in \mathbb{C}\mathbf{1}$ (Exercise 4.3.6), so that $\mathrm{tr}\, x = 0$ leads to $x = 0$. Hence ad is injective, and we conclude with $\dim \mathfrak{so}(E,\beta) = \dim \mathfrak{so}_3(\mathbb{R}) = 3$ that

$$\mathrm{ad}(\mathfrak{su}_2(\mathbb{C})) = \mathfrak{so}(E,\beta)$$

(cf. Exercise 4.3.7). Therefore the connectedness of $\mathrm{SU}_2(\mathbb{C})$ implies that

$$\mathrm{Ad}(\mathrm{SU}_2(\mathbb{C})) = \langle \mathrm{Ad}(\exp \mathfrak{su}_2(\mathbb{C}))\rangle = \langle e^{\mathrm{ad}\, \mathfrak{su}_2(\mathbb{C})}\rangle = \langle e^{\mathfrak{so}(E,\beta)}\rangle = \mathrm{SO}(E,\beta)_0 = \mathrm{SO}(E,\beta)$$

(cf. Exercise 4.3.8). We thus obtain a surjective homomorphism

$$\varphi\colon \mathrm{SU}_2(\mathbb{C}) \to \mathrm{SO}_3(\mathbb{R}).$$

Since $\mathrm{SU}_2(\mathbb{C})$ is compact, the quotient group $\mathrm{SU}_2(\mathbb{C})/\ker\varphi$ is also compact, and the induced bijective morphism $\mathrm{SU}_2(\mathbb{C})/\ker\varphi \to \mathrm{SO}_3(\mathbb{R})$ is a homeomorphism, hence an isomorphism of topological groups.

We further have

$$\ker\varphi = Z(\mathrm{SU}_2(\mathbb{C})) = \mathrm{SU}_2(\mathbb{C}) \cap \mathbb{C}^\times\mathbf{1} = \{\pm\mathbf{1}\}$$

(Exercise 1.1.9), so that

$$\widetilde{\mathrm{SO}_3}(\mathbb{R}) \cong \mathrm{SU}_2(\mathbb{C}) \quad \text{and} \quad \pi_1(\mathrm{SO}_3(\mathbb{R})) \cong C_2.$$

Exercises for Section 4.3

Exercise 4.3.1. Let G be an abelian group and $N \leq G$ a subgroup carrying a Lie group structure. Then there exists a unique Lie group structure on G for which N is an open subgroup.

Exercise 4.3.2. Let G be a Lie group and $N \trianglelefteq G$ a discrete normal subgroup. Show that G/N carries a unique Lie group structure for which the quotient map $q\colon G \to G/N$ is a local diffeomorphism.

Exercise 4.3.3. Let M, N and X be smooth manifolds, $q\colon M \to N$ be a smooth covering map and $F\colon X \to M$ a continuous map. Show that F is smooth if and only if $q \circ F$ is smooth.

Exercise 4.3.4. Let $q_G\colon \widetilde{G} \to G$ be a simply connected covering of the connected Lie group G.

(1) Show that each automorphism $\varphi \in \mathrm{Aut}(G)$ has a unique lift $\widetilde{\varphi} \in \mathrm{Aut}(\widetilde{G})$.

(2) Derive from (1) that $\mathrm{Aut}(G) \cong \{\widetilde{\varphi} \in \mathrm{Aut}(\widetilde{G}) \colon \widetilde{\varphi}(\pi_1(G)) = \pi_1(G)\}$.

(3) Show that, in general, $\{\widetilde{\varphi} \in \mathrm{Aut}(\widetilde{G}) \colon \widetilde{\varphi}(\pi_1(G)) \subseteq \pi_1(G)\}$ is not a subgroup of $\mathrm{Aut}(\widetilde{G})$.

Exercise 4.3.5. Let $(E, +)$ be the additive group of a real Banach space, considered as a Lie group. Show that $\mathrm{Aut}(E, +) = \mathrm{GL}(E)$, i.e., every automorphism of $(E, +)$ is linear.

Exercise 4.3.6. Show that the center of $\mathrm{GL}_n(\mathbb{K})$ is given by

$$Z(\mathrm{GL}_n(\mathbb{K})) := \{g \in \mathrm{GL}_n(\mathbb{K}) \colon (\forall h \in \mathrm{GL}_n(\mathbb{K}))\, hg = gh\} = \mathbb{K}^\times \mathbf{1}$$

and that the center of its Lie algebra is

$$\mathfrak{z}(\mathfrak{gl}_n(\mathbb{K})) := \{x \in \mathfrak{gl}_n(\mathbb{K}) \colon (\forall x \in \mathfrak{gl}_n(\mathbb{K}))\, [x, y] = 0\} = \mathbb{K}\mathbf{1}.$$

Hint: Consider the elementary matrices $E_{ij} := (\delta_{ik}\delta_{jl})_{k,l}$ and note that $T_{ij} := \mathbf{1} + E_{ij} \in \mathrm{GL}_n(\mathbb{K})$.

Exercise 4.3.7. Let (E, β) be an n-dimensional euclidean space, i.e., β is a positive definite symmetric bilinear form on E. Show that there exists an isometric isomorphism $\Phi \colon \mathbb{R}^n \to E$, and that

$$\Psi \colon \mathrm{O}_n(\mathbb{R}) \to \mathrm{O}(E, \beta), \quad g \mapsto \Phi \circ g \circ \Phi^{-1}$$

is an isomorphism of topological groups.

Exercise 4.3.8. Show that for every linear Lie group $G \subseteq \mathcal{A}^\times$, the identity component is generated by the image of $\exp_G$:

$$G_0 = \langle \exp_G(\mathbf{L}(G)) \rangle.$$

Exercise 4.3.9. On the four-dimensional real vector space $V := \mathrm{Herm}_2(\mathbb{C})$ we consider the symmetric bilinear form β given by

$$\beta(A, B) := \mathrm{tr}(AB) - \mathrm{tr}\, A \,\mathrm{tr}\, B.$$

Show that:

(1) The corresponding quadratic form is given by $q(A) := \beta(A, A) = -2 \det A$.

(2) Show that $(V, \beta) \cong \mathbb{R}^{3,1}$ by finding a basis $E_1, \ldots, E_4$ of $\mathrm{Herm}_2(\mathbb{C})$ with

$$q(a_1 E_1 + \ldots + a_4 E_4) = a_1^2 + a_2^2 + a_3^2 - a_4^2.$$

(3) For $g \in \mathrm{GL}_2(\mathbb{C})$ and $A \in \mathrm{Herm}_2(\mathbb{C})$ the matrix gAg^* is hermitian and satisfies

$$q(gAg^*) = |\det(g)|^2 q(A).$$

(4) For $g \in \mathrm{SL}_2(\mathbb{C})$ we define a linear map $\rho(g) \in \mathrm{GL}(\mathrm{Herm}_2(\mathbb{C}))$ by $\rho(g)(A) := gAg^*$. Then we obtain a homomorphism

$$\rho\colon \mathrm{SL}_2(\mathbb{C}) \to \mathrm{O}(V, \beta) \cong \mathrm{O}_{3,1}(\mathbb{R}).$$

(5) Show that $\ker \rho = \{\pm \mathbf{1}\}$.

Exercise 4.3.10. Let $\varphi\colon G \to H$ be a surjective morphism of topological groups. Show that the following conditions are equivalent:

(1) φ is open with discrete kernel.

(2) φ is a covering in the topological sense, i.e., each $h \in H$ has an open neighborhood U such that $\varphi^{-1}(U) = \bigcup_{i \in I} U_i$ is a disjoint union of open subsets U_i for which all restrictions $\varphi|_{U_i} : U_i \to U$ are homeomorphisms.

Chapter 5

Vector Fields and their Lie-Algebra Structure

At this point we have seen several constructions producing new Lie groups from given ones. In particular, we have seen that linear Lie groups are Lie groups and that coverings of Lie groups, resp., quotients by discrete normal subgroups are Lie groups. What is still missing is a definition of the Lie algebra of a general Lie group, i.e., the corresponding infinitesimal object.

The idea to define $\mathbf{L}(G)$ is the following. For linear Lie groups $G \subseteq \mathcal{A}^\times$, we have defined $\mathbf{L}(G)$ as those elements of $\mathcal{A}$ for which $e^{\mathbb{R}x} \subseteq G$, and it is easy to see that, geometrically, $\mathbf{L}(G)$ is the tangent space of G in $\mathbf{1}$ (Exercise 3.1.8). Therefore each $x \in \mathbf{L}(G)$ defines by $X(g) = gx$ a "vector field" on G. For general Lie groups, this is a natural way to obtain the Lie algebra as the set of vector fields on G which are invariant under left translations. In this chapter we provide the concepts needed to follow this path. First we define tangent vectors of abstract manifolds (Section 5.1). Then we define vector fields and show that the space of vector fields carries a natural Lie bracket (Section 5.2), and finally we explain how vector fields are related to flows, resp., ordinary differential equations on manifolds (Section 5.3).

5.1 Tangent vectors of manifolds

The real strength of the theory of smooth manifolds is due to the fact that it permits to analyze differentiable structures in terms of their derivatives. To model these derivatives appropriately, we introduce the tangent bundle TM of a smooth manifold, tangent maps of smooth maps and smooth vector fields.

Remark 5.1.1. To understand the idea behind tangent vectors and the tangent bundle of a manifold, we first take a closer look at the special case of an open subset U of a Banach space E. We think of a *tangent vector in* $p \in U$ as a vector $v \in E$ attached to the point p. In particular, we distinguish tangent vectors attached to different points. A good way to visualize this is to think of v as an arrow starting in p or as v

corresponding to the translation map $\tau_v \colon E \to E, x \mapsto x + v$. In this sense we write

$$T_p(U) := \{p\} \times E$$

for the set of all tangent vectors in p, the *tangent space in* p. It carries a natural real vector space structure, given by

$$(p, v) + (p, w) := (p, v + w) \quad \text{and} \quad \lambda(p, v) := (p, \lambda v)$$

for $v, w \in E$ and $\lambda \in \mathbb{R}$.

The collection of all tangent vectors, the *tangent bundle*, is denoted

$$T(U) := \bigcup_{p \in U} T_p(U) = \{(p, v) \colon p \in U, v \in E\} \cong U \times E.$$

(b) Let E and F be Banach spaces. If $f \colon U \to V$ is a C^1-map between open subsets $U \subseteq E$ and $V \subseteq F$, then the directional derivatives permit us to "extend" f to tangent vectors by its *tangent map*

$$T(f) \colon T(U) \cong U \times E \to TV \cong V \times F, \quad (p, v) \mapsto (f(p), \mathtt{d}f(p)v).$$

For each $p \in U$, the map Tf restricts to a linear map

$$T_p(f) \colon T_p(U) \to T_{f(p)}(V), \quad (p, v) \mapsto (f(p), \mathtt{d}f(p)v). \tag{5.1}$$

The main difference to the map $\mathtt{d}f$ is the book keeping; here we keep track of what happens to the point p and the tangent vector v.

If $U \subseteq E$, $V \subseteq F$, resp., $W \subseteq G$ are open subsets, where E, F and G are Banach spaces, and $f \colon U \to V$ and $g \colon V \to W$ are C^1-maps, then the Chain Rule implies that

$$\begin{aligned} T(g \circ f)(p, v) &= (g(f(p)), \mathtt{d}(g \circ f)(p)v) = (g(f(p)), \mathtt{d}g(f(p))\mathtt{d}f(p)v) \\ &= T_{f(p)}(g) \circ T_p(f)v, \end{aligned}$$

so that, in terms of tangent maps, the Chain Rule takes the simple form

$$T(g \circ f) = T(g) \circ T(f). \tag{5.2}$$

We clearly have $T(\mathrm{id}_U) = \mathrm{id}_{T(U)}$ and, for a diffeomorphism f, we thus obtain from the Chain Rule $T(f^{-1}) = T(f)^{-1}$.

(c) As the terminology suggests, tangent vectors arise as tangent vectors of curves. If $\gamma \colon]a, b[\to U$ is a C^1-curve, then its tangent map satisfies

$$T(\gamma)(t, v) = (\gamma(t), \mathtt{d}\gamma(t)v) = (\gamma(t), v \cdot \gamma'(t)) \quad \text{on} \quad T(]a, b[) \cong]a, b[\times \mathbb{R},$$

and in particular

$$T(\gamma)(t, 1) = (\gamma(t), \gamma'(t))$$

is the *tangent vector in* $\gamma(t)$ to the curve γ.

Definition 5.1.2. Let M be a smooth E-manifold and $\mathcal{A} := (\varphi_i, U_i)_{i\in I}$ be an E-atlas of M. On the disjoint union of the set $\varphi(U_i) \times E$, we define an equivalence relation by

$$(\varphi_i(p), v) \sim \big(\varphi_j(p), \mathtt{d}(\varphi_j \circ \varphi_i^{-1})(\varphi_i(p))(v)\big) = T(\varphi_j \circ \varphi_i^{-1})(\varphi_i(p), v),$$

for $p \in U_i \cap U_j$ and $v \in E$.[1] We write $[\varphi_i(p), v]$ for the equivalence class of $(\varphi_i(p), v)$. Then the equivalence classes of the form $[\varphi_i(p), v]$, $v \in E$, are called *tangent vectors in* p and we write $T_p(M)$ for the set of all these equivalence classes, the *tangent space of* M *in* p. Since the differentials $\mathtt{d}(\varphi_j \circ \varphi_i^{-1})(\varphi_i(p))$ are invertible linear maps, $T_p(M)$ inherits the well-defined structure of a vector space by

$$[\varphi_i(p), v] + [\varphi_i(p), w] := [\varphi_i(p), v + w] \quad \text{ and } \quad \lambda[\varphi_i(p), v] := [\varphi_i(p), \lambda v],$$

so that the map $E \to T_p(M), v \mapsto [\varphi_i(p), v]$ is a linear isomorphism for any fixed $i \in I$ (Exercise).

Remark 5.1.3. If (φ, U) is a chart of M compatible with the E-atlas $\mathcal{A}$ and $p \in U$, $v \in E$, then the equivalence class $[\varphi(p), v]$ is defined, regardless of whether (φ, U) was contained in the atlas $\mathcal{A}$ or not. In fact, we can always enlarge $\mathcal{A}$ by the chart (φ, U) and observe that, for any $i \in I$ with $p \in U_i$, we have

$$[\varphi(p), v] = [\varphi_i(p), \mathtt{d}(\varphi_i \circ \varphi^{-1})(\varphi(p))v],$$

expressing $[\varphi(p), v]$ in terms of the equivalence classes specified by the atlas $\mathcal{A}$ in Definition 5.1.2.

Definition 5.1.4. (The tangent bundle TM) As a set, the *tangent bundle* of the E-manifold M is defined as the disjoint union

$$TM := \bigcup_{p\in M} T_p(M)$$

of all tangent vectors. To define a topology, resp., a manifold structure on M, we first recall that any E-chart (φ, U) of M leads to a bijection

$$T\varphi\colon TU = \bigcup_{p\in U} T_p(M) \to T(\varphi(U)) = \varphi(U) \times E, \quad [\varphi(p), v] \mapsto (\varphi(p), v).$$

Let $\tau \subseteq \mathbb{P}(TM)$ be the set of all subsets $O \subseteq TM$ with the property that $(T\varphi)(O)$ is open in $\varphi(U) \times E$ for every E-chart (φ, U) of M. It is easy to see that τ is a topology TM. Since for every open subset $O' \subseteq \varphi(U) \times E$ for some chart (φ, U) of M, the sets $T(\psi \circ \varphi^{-1})(O')$ are open in $\psi(V) \times E$ for every other chart (ψ, V) of M, it follows that $(T\varphi)^{-1}(O') \in \tau$, and therefore $T\varphi\colon TU \to \varphi(U) \times E$ is a homeomorphism, hence an $(E \times E)$-chart of TM.

Let

$$\pi_{TM}\colon TM \to M, \quad T_p(M) \ni v \mapsto p$$

[1]That this defines an equivalence relation follows easily from the fact that the maps $\varphi_j \circ \varphi_i^{-1}$ are diffeomorphisms and the Chain Rule.

denote the projection map sending a tangent vector $v \in T_p(M)$ to its base point p. Then, for every open subset $O \subseteq M$ its inverse image $\pi_{TM}^{-1}(O) \subseteq TM$ is open because for every chart (φ, U), the set $\varphi(O \cap U) \times E$ is open in $E \times E$. To see that TM is Hausdorff, we now observe that, for two tangent vectors $v \in T_p(M)$ and $w \in T_q(M)$, we obtain disjoint open neighborhoods from the continuity of π_{TM} if $p \neq q$. If $p = q$, we pick any chart (φ, U) with $p \in U$ and note that the open subset $TU \cong \varphi(U) \times E$ of TM is Hausdorff, so that $v, w \in TU$ possess disjoint open neighborhoods.

Finally we observe that, for every E-atlas $\mathcal{A} = (\varphi_i, U_i)_{i \in I}$, we obtain an $(E \times E)$-atlas $T\mathcal{A} := (T\varphi_i, TU_i)_{i \in I}$ (Exercise 5.1.1).

Remark 5.1.5. For each open subset U of a Banach space E, we have $TU \cong U \times E$ (as smooth manifolds).

Definition 5.1.6. (The tangent map) Let M and N be smooth manifolds and $f\colon M \to N$ be a smooth map.

(a) For $p \in M$ we pick charts (φ, U) of M and (ψ, V) of N with $p \in U$ and $f(p) \in V$. We now define the *tangent map*

$$T_p(f) := \mathtt{d}f(p)\colon T_p(M) \to T_{f(p)}(N), \quad [\varphi(p), v] \mapsto [\psi(f(p)), \mathtt{d}(\psi \circ f \circ \varphi^{-1})(\varphi(p))v].$$

The Chain Rule implies that the right hand side does not depend on the choice of the charts of M, resp., N. We thus obtain a well-defined linear map $T_p(f)$ between the tangent spaces.

Putting all these tangent maps together, we obtain a map

$$Tf\colon TM \to TN, \quad Tf(v) := T_p(f)v, \quad v \in T_p(M).$$

For charts (φ, U) on M and (ψ, V) as above, we have

$$T\psi \circ Tf \circ (T\varphi)^{-1}(\varphi(p), v) = (\psi(f(p)), \mathtt{d}(\psi \circ f \circ \varphi^{-1})(\varphi(p))v) = T(\psi \circ f \circ \varphi^{-1})(\varphi(p), v),$$

so that Tf is smooth on the open subset $T(f^{-1}(V))$ of TM. This implies that Tf is a smooth map.

(b) If $N = E$ is a Banach space, then $TE \cong E \times E$ and the map Tf can accordingly be written as $Tf = (f, \mathtt{d}f)$, where $\mathtt{d}f$ is a map $TM \to E$.

(c) If $M \subseteq \mathbb{R}$ is an open interval, then f is a curve in M and we put

$$f'(t) := (T_t f)(1) \in T_{f(t)}(N).$$

Remark 5.1.7. (a) Note that

$$T(\mathrm{id}_M) = \mathrm{id}_{TM} \quad \text{and} \quad T(f_1 \circ f_2) = Tf_1 \circ Tf_2$$

for smooth maps $f_2\colon M_1 \to M_2$ and $f_2\colon M_2 \to M_3$.

(b) For smooth manifolds $M_1, \ldots, M_n$, the projection maps

$$\pi_i\colon M_1 \times \cdots \times M_n \to M_i, \quad (p_1, \ldots, p_n) \mapsto p_i$$

induce a diffeomorphism

$$(T\pi_1, \ldots, T\pi_n)\colon T(M_1 \times \cdots \times M_n) \to TM_1 \times \cdots \times TM_n.$$

Remark 5.1.8. We record some rules that are useful when dealing with tangent maps.

(1) From the local Chain Rule ((5.2) in Remark 5.1.1) we immediately derive its general form: For smooth maps $g\colon M_1 \to M_2$ and $f\colon M_2 \to M_3$, we have

$$T(f \circ g) = T(f) \circ T(g). \tag{5.3}$$

We also note that $T(\mathrm{id}_M) = \mathrm{id}_{TM}$. [2]

(2) If $f_1\colon M \to N_1$ and $f_2\colon M \to N_2$ are smooth maps, then we combine them to the smooth map

$$f := (f_1, f_2)\colon M \to N_1 \times N_2, \quad m \mapsto (f_1(m), f_2(m)).$$

Identifying $T(N_1 \times N_2)$ with $TN_1 \times TN_2$, we then have

$$Tf = (Tf_1, Tf_2).$$

(3) If $f_1\colon M_1 \to N_1$ and $f_2\colon M_2 \to N_2$ are smooth maps, then we combine them to the smooth map

$$f := f_1 \times f_2\colon M_1 \times M_2 \to N_1 \times N_2, \quad (m_1, m_2) \mapsto (f_1(m_1), f_2(m_2)).$$

Identifying $T(M_1 \times M_2)$ with $TM_1 \times TM_2$ and $T(N_1 \times N_2)$ with $TN_1 \times TN_2$, we obtain

$$T(f_1 \times f_2) = Tf_1 \times Tf_2.$$

Exercises for Section 5.1

Exercise 5.1.1. Let $(\varphi_i, U_i)_{i\in I}$ be an E-chart of the smooth manifold M and

$$\psi_i\colon TU_i := \bigcup_{p\in U_i} T_p(M) \to \varphi(U_i) \times E, \quad [\varphi_i(x), v] \mapsto (\varphi_i(x), v).$$

Show that

(i) We call $O \subseteq TM$ *open* if, for each $i \in I$, the subset $\psi_i(O \cap TU_i) \subseteq \varphi_i(U_i) \times E$ is open. Show that the open sets define a topology on TM.

(ii) The subset $TU_i \subseteq TM$ is open and $\psi_i\colon TU_i \to \varphi(U_i) \times E$ is a homeomorphism.

(iii) TM is a Hausdorff space.

(iv) The family $(\psi_i)_{i\in I}$ defines a smooth $E \times E$-atlas on TM.

Exercise 5.1.2. Let M be an E-manifold. Show that, for every E-chart (φ, U) of M, the tangent map

$$T\varphi\colon TU \to T(\varphi(U)) \cong \varphi(U) \times E$$

is a diffeomorphism.

[2] Both observations combine to the insight that the assignments $M \mapsto TM$ and $f \mapsto Tf$ define a functor from the category of smooth manifolds into itself.

Exercise 5.1.3. [Inverse Function Theorem for manifolds] Let $f\colon M \to N$ be a smooth map between manifold and $p \in M$ such that $T_p(f)\colon T_p(M) \to T_p(N)$ is a linear isomorphism. Show that there exists an open neighborhood U of p in M such that the restriction $f|_U\colon U \to f(U)$ is a diffeomorphism onto an open subset of N.

Exercise 5.1.4. If $f\colon M \to N$ is a smooth map and $\pi_{TM}\colon TM \to M$ is the projection map of the tangent bundle, then

$$\pi_{TN} \circ Tf = f \circ \pi_{TM}.$$

5.2 The Lie algebra of vector fields

Definition 5.2.1. We define a *(smooth) vector field* X on M as a smooth map $X\colon M \to TM$ with $\pi_{TM} \circ X = \mathrm{id}_M$, where $\pi_{TM}\colon TM \to M$ denotes the projection map mapping $T_p(M)$ to p. Alternatively, we can say that $X(p) \in T_p(M)$ for every $p \in M$. We write $\mathcal{V}(M)$ for the space of all vector fields on M.

If $f \in C^\infty(M, F)$ is a smooth function on M with values in some Banach space F and $X \in \mathcal{V}(M)$, then we obtain a smooth function on M via

$$Xf := \mathtt{d}f \circ X\colon M \to TM \to F.$$

In each point $p \in M$, we then have

$$(Xf)(p) = \mathtt{d}f(p)X(p),$$

so that $(Xf)(p)$ is a directional derivative of the function f in the direction of the tangent vector $X(p)$.

Remark 5.2.2. If $M = U$ is an open subset of the Banach space E, then $TU = U \times E$, with the projection $\pi_{TU}\colon U \times E \to U, (x, v) \mapsto x$. Therefore each smooth vector field is of the form $X(x) = (x, \widetilde{X}(x))$ for some smooth function $\widetilde{X}\colon U \to E$. In this sense we may thus identify $\mathcal{V}(U)$ with the space $C^\infty(U, E)$.

Definition 5.2.3. If $\varphi\colon M \to N$ is a smooth map, then we say that $X \in \mathcal{V}(M)$ and $Y \in \mathcal{V}(N)$ are *φ-related* if

$$T\varphi \circ X = Y \circ \varphi\colon M \to TN. \tag{5.4}$$

If, in addition, φ is a diffeomorphism, then this implies that

$$Y = T\varphi \circ X \circ \varphi^{-1} =: \varphi_* X.$$

For a vector field $X \in \mathcal{V}(M)$, we call $\varphi_* X \in \mathcal{V}(N)$ the corresponding φ-transformed vector field on N. It is the unique vector field on N which is φ-related to X.

Remark 5.2.4. Suppose that $\varphi\colon M \to N$ is a smooth map and $f\colon N \to F$ is a smooth map with values in the Banach space F. If $X \in \mathcal{V}(M)$ and $Y \in \mathcal{V}(N)$ are φ-related, then

$$X(f \circ \varphi) = (Yf) \circ \varphi \tag{5.5}$$

follows from

$$X(f \circ \varphi) = \mathtt{d}(f \circ \varphi) \circ X = \mathtt{d}f \circ T\varphi \circ X = \mathtt{d}f \circ Y \circ \varphi = (Yf) \circ \varphi.$$

Lemma 5.2.5. *Let $U \subseteq E$ be an open subset of the Banach space E and identify vector fields on U with smooth E-valued functions. Then we obtain a Lie bracket on $C^\infty(U, E)$ by*

$$[X, Y](p) := \mathtt{d}Y(p)X(p) - \mathtt{d}X(p)Y(p) \quad \textit{for} \quad p \in U. \tag{5.6}$$

With respect to this Lie bracket, the map

$$\mathcal{L} \colon \mathcal{V}(U) \cong C^\infty(U, E) \to \operatorname{End}(C^\infty(U, E)), \quad X \mapsto \mathcal{L}_X, \quad \mathcal{L}_X(f)(p) := \mathtt{d}f(p)X(p)$$

is an injective homomorphism of Lie algebras, in particular,

$$\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]. \tag{5.7}$$

Proof. Let $V := C^\infty(U, E)$ and consider $\mathcal{L}_X$ as a linear endomorphism of V. If $\mathcal{L}_X = 0$, then $0 = \mathcal{L}_X \operatorname{id}_U = X$, so that $\mathcal{L}$ is injective. For $f \in V$ we obtain

$$\mathcal{L}_X\mathcal{L}_Y(f)(p) = \mathtt{d}(\mathcal{L}_Y f)(p)X(p) = (\mathtt{d}^2 f)(p)(X(p), Y(p)) + \mathtt{d}f(p)\mathtt{d}Y(p)X(p),$$

so that the Schwarz Lemma, i.e., the symmetry of the second differential,[3] implies

$$\mathcal{L}_X\mathcal{L}_Y f - \mathcal{L}_Y\mathcal{L}_X f = \mathcal{L}_{[X,Y]} f,$$

which is (5.7).

Clearly, the bracket on $\mathcal{V}(U)$ is skew-symmetric. That it also satisfies the Jacobi identity follows from the injectivity of $\mathcal{L}$ and the Jacobi identity in $\operatorname{End}(V)$:

$$\mathcal{L}_{[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]} = [\mathcal{L}_X, [\mathcal{L}_Y, \mathcal{L}_Z]] + [\mathcal{L}_Y, [\mathcal{L}_Z, \mathcal{L}_X]] + [\mathcal{L}_Z, [\mathcal{L}_X, \mathcal{L}_Y]]. \quad \square$$

Having dealt with the local version of the Lie bracket on vector fields, we now turn to its global form.

Lemma 5.2.6. *If $X, Y \in \mathcal{V}(M)$, then there exists a vector field $[X, Y] \in \mathcal{V}(M)$ which is uniquely determined by the property that, for each Banach space F and each open subset $U \subseteq M$, we have*

$$[X, Y]f = X(Yf) - Y(Xf) \quad \textit{for} \quad f \in C^\infty(U, F). \tag{5.8}$$

[3] To formulate the Schwarz Lemma for a C^2-map $f \colon U \to F$ from the open subset U of the Banach space E to the Banach space F, we define the *second differential of f* as the map $\mathtt{d}^2 f \colon U \to \mathcal{L}^2(E^2, F)$ by

$$\mathtt{d}^2 f(p)(h_1, h_2) = \frac{d}{dt}\Big|_{t=0} \mathtt{d}f(p + th_2)(h_1).$$

Then the Schwarz Lemma asserts that all the bilinear maps $\mathtt{d}^2 f(p)$ are symmetric. It follows from the special case $E = \mathbb{R}^2$, applied to the map $(t_1, t_2) \mapsto f(p + t_1 h_1 + t_2 h_2)$. One should also observe that it follows from the formula

$$\mathtt{d}^2 f(p)(h_1, h_2) = \lim_{n\to\infty} \frac{1}{n^2}\Big(f\Big(p + \frac{1}{n}h_1 + \frac{1}{n}h_2\Big) - f\Big(p + \frac{1}{n}h_1\Big) - f\Big(p + \frac{1}{n}h_2\Big) + f(p)\Big).$$

Proof. **Uniqueness:** First we show that $[X,Y]$ is uniquely determined by (5.8). If (φ, U) is an E-chart of M, then

$$[X,Y]\varphi = \mathtt{d}\varphi \circ [X,Y] \colon U \to E$$

determines $[X,Y]$ because the linear maps $\mathtt{d}\varphi(p), p \in U$, are injective.

Existence: Let (φ, U) be an E-chart of M. Then $A := \varphi_* X := T\varphi \circ X \circ \varphi^{-1}$ and $B := \varphi_* Y := T\varphi \circ Y \circ \varphi^{-1}$ are smooth vector fields on the open subset $\varphi(U) \subseteq E$.

For a smooth function $f \colon U' \to F$ on an open subset $U' \subseteq U$ we obtain a smooth function $h := f \circ \varphi^{-1} \colon \varphi(U') \to F$, and

$$Xf \circ \varphi^{-1} = \mathtt{d}f \circ X \circ \varphi^{-1} = \mathtt{d}f \circ T(\varphi^{-1}) \circ T(\varphi) \circ X \circ \varphi^{-1} = \mathtt{d}h \circ A = Ah.$$

Representing the vector fields A and B by smooth functions

$$A(p) = \big(p, \widetilde{A}(p)\big), \quad B(p) = \big(p, \widetilde{B}(p)\big),$$

we have

$$Ah(p) = \mathtt{d}h(p)\widetilde{A}(p) \quad \text{and} \quad Bh(p) = \mathtt{d}h(p)\widetilde{B}(p),$$

so that Lemma 5.2.5 implies that

$$ABh - BAh = Ch$$

holds for the vector field $C(p) = (p, \widetilde{C}(p))$ with

$$\widetilde{C}(p) := \mathtt{d}\widetilde{B}(p)\big(\widetilde{A}(p)\big) - \mathtt{d}\widetilde{A}(p)\big(\widetilde{B}(p)\big). \tag{5.9}$$

The smoothness of the right hand side follows from the chain rule. For the smooth vector field

$$D_U := (\varphi^{-1})_* C = T\varphi^{-1} \circ C \circ \varphi \in \mathcal{V}(U),$$

we obtain with (5.5)

$$\begin{aligned} D_U f &= D_U(h \circ \varphi) = Ch = A(Bh) - B(Ah) \\ &= X(Bh \circ \varphi^{-1}) - B(Ah \circ \varphi^{-1}) = X(Yf) - Y(Xf). \end{aligned}$$

This implies the existence of a vector field $[X,Y]_U := D_U$ satisfying (5.8) on the open subset U.

From the uniqueness that we have already verified, we obtain

$$[X,Y]_{U\cap U'} = [X,Y]_U|_{U\cap U'} = [X,Y]_{U'}|_{U\cap U'}$$

for two open subsets $U, U' \subseteq M$. Therefore $[X,Y](p) := [X,Y]_U(p)$ for $p \in U$, yields a well-defined vector field $[X,Y] \in \mathcal{V}(M)$, satisfying (5.8) on M. □

Proposition 5.2.7. $(\mathcal{V}(M), [\cdot,\cdot])$ *is a Lie algebra.*

Proof. Since the Lie bracket is obviously skew symmetric, we have to verify the Jacobi identity, i.e., the vanishing of all vector fields of the form

$$J(X,Y,Z) := [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]].$$

It suffices to show for every chart (φ, U) of M that $J(X,Y,Z)|_U = 0$. This follows from the fact that the construction of the bracket on $\mathcal{V}(M)$ shows that

$$\Phi\colon \mathcal{V}(U) \to \mathcal{V}(\varphi(U)), \quad X \mapsto \varphi_* X$$

is compatible with the bracket in the sense that $\Phi([X,Y]) = [\Phi(X), \Phi(Y)]$ (see the proof of Lemma 5.2.6). Therefore the Jacobi identity in $\mathcal{V}(U)$ follows from the Jacobi identity on $\mathcal{V}(\varphi(U))$ (Lemma 5.2.5). □

For the applications to Lie groups we will need the following lemma.

Lemma 5.2.8. (Related Vector Field Lemma) *Let M and N be smooth manifolds, $\varphi\colon M \to N$ be a smooth map, and $X_M, Y_M \in \mathcal{V}(M)$, $X_N, Y_N \in \mathcal{V}(N)$ such that the pairs (X_M, X_N) and (Y_M, Y_N) are φ-related. Then their Lie brackets $[X_M, Y_M]$ and $[X_N, Y_N]$ are also φ-related.*

Proof. For every chart (ψ, V) of N we obtain with (5.5)

$$\begin{aligned}([X_N, Y_N]\psi) \circ \varphi &= X_N(Y_N\psi) \circ \varphi - Y_N(X_N\psi) \circ \varphi = X_M(Y_N\psi \circ \varphi) - Y_M(X_N\psi \circ \varphi) \\ &= X_M(Y_M(\psi \circ \varphi)) - Y_M(X_N(\psi \circ \varphi)) = [X_M, Y_M](\psi \circ \varphi)\end{aligned}$$

For every $p \in M$, this implies that

$$\mathtt{d}\psi(\varphi(p))[X_N, Y_N](\varphi(p)) = \mathtt{d}\psi(\varphi(p)) T_p(\varphi)[X_M, Y_M](p),$$

and since $\mathtt{d}\psi(p)$ is injective, we see that

$$[X_N, Y_N] \circ \varphi = T\varphi \circ [X_M, Y_M],$$

which completes the proof. □

Exercises for Section 5.2

Exercise 5.2.1. Let $\mathcal{A}$ be a $\mathbb{K}$-algebra (not necessarily associative). Show that

(i) $\mathrm{der}(\mathcal{A}) := \{D \in \mathrm{End}(\mathcal{A})\colon (\forall a, b \in \mathcal{A})\, D(ab) = Da \cdot b + a \cdot Db\}$ is a Lie subalgebra of $\mathfrak{gl}(\mathcal{A}) = \mathrm{End}(\mathcal{A})_L$.

(ii) If, in addition, $\mathcal{A}$ is commutative, then for $D \in \mathrm{der}(\mathcal{A})$ and $a \in \mathcal{A}$, the map $aD\colon \mathcal{A} \to \mathcal{A}, x \mapsto aDx$ also is a derivation.

Exercise 5.2.2. Let U be an open subset of $\mathbb{R}^{2n}$ and $\mathcal{P} = C^\infty(U, \mathbb{R})$ be the set of smooth functions on U and write $q_1, \ldots, q_m, p_1, \ldots, p_m$ for the coordinates with respect to a basis. Then $\mathfrak{g}$ is a Lie algebra with respect to the *Poisson bracket*

$$\{f, g\} := \sum_{i=1}^{n} \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}.$$

Exercise 5.2.3. Let U be an open subset of $\mathbb{R}^n$, $\mathcal{A} = C^\infty(U,\mathbb{R})$, and $\mathcal{V} = C^\infty(U,\mathbb{R}^n)$. For $f \in A$ and $X = (X_1,\ldots,X_n) \in \mathcal{V}$, we define

$$\mathcal{L}_X f := Xf := \sum_{i=1}^n X_i \frac{\partial f}{\partial x_i}.$$

(i) The maps $\mathcal{L}_X$ are derivations of the algebra $\mathcal{A}$.

(ii) If $\mathcal{L}_X = 0$, then $X = 0$.

(iii) The commutator of two such operators has the form $[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_{[X,Y]}$, where the bracket on $\mathcal{V}$ is defined by

$$[X,Y](p) := \mathtt{d}Y(p)X(p) - \mathtt{d}X(p)Y(p),$$

resp.,

$$[X,Y]_i = \sum_{j=1}^n X_j \frac{\partial Y_i}{\partial x_j} - Y_j \frac{\partial X_i}{\partial x_j}.$$

(iv) $(\mathcal{V}, [\cdot,\cdot])$ is a Lie algebra.

(v) To each $A \in \mathfrak{gl}_n(\mathbb{R})$, we associate the linear vector field $X_A(x) := Ax$. Show that, for $A, B \in M_n(\mathbb{R})$, we have $X_{[A,B]} = -[X_A, X_B]$.

5.3 Flows and vector fields

In this section we turn to the geometric nature of vector fields as infinitesimal generators of flows on manifolds. This provides a natural global perspective on ordinary differential equations.

Throughout this section, M denotes an E-manifold for some Banach space E.

Definition 5.3.1. Let $X \in \mathcal{V}(M)$ and $I \subseteq \mathbb{R}$ be an open interval containing 0. A differentiable map $\gamma\colon I \to M$ is called an *integral curve* of X if

$$\gamma'(t) = X(\gamma(t)) \qquad \text{for each} \quad t \in I.$$

Note that the preceding equation implies that γ' is continuous and further that if γ is C^k, then γ' is also C^k. Therefore integral curves are automatically smooth.

If $J \supseteq I$ is an interval containing I, then an integral curve $\eta\colon J \to M$ is called an *extension* of γ if $\eta|_I = \gamma$. An integral curve γ is said to be *maximal* if it has no proper extension.

Remark 5.3.2. (a) If $U \subseteq E$ is an open subset of the Banach space E, then we write a vector field $X \in \mathcal{V}(U)$ as $X(x) = (x, F(x))$, where $F\colon U \to E$ is a smooth function. A curve $\gamma\colon I \to U$ is an integral curve of X if and only if it satisfies the ordinary differential equation

$$\gamma'(t) = F(\gamma(t)) \quad \text{for all} \quad t \in I.$$

(b) If (φ, U) is a chart of the manifold M and $X \in \mathcal{V}(M)$, then a curve $\gamma\colon I \to M$ is an integral curve of X if and only if the curve $\eta := \varphi \circ \gamma$ is an integral curve of the vector field $X_\varphi := \varphi_* X = T(\varphi) \circ X \circ \varphi^{-1} \in \mathcal{V}(\varphi(U))$ because

$$X_\varphi(\eta(t)) = T_{\gamma(t)}(\varphi)X(\gamma(t)) \quad \text{and} \quad \eta'(t) = T_{\gamma(t)}(\varphi)\gamma'(t).$$

Remark 5.3.3. A curve $\gamma\colon I \to M$ is an integral curve of X if and only if $\widetilde{\gamma}(t) := \gamma(-t)$ is an integral curve of the vector field $-X$.

More generally, for $a, b \in \mathbb{R}$, the curve $\eta(t) := \gamma(at + b)$ is an integral curve of the vector field aX.

For the proof of the following theorem we refer to [La99, Thm. 4.2.1].

Theorem 5.3.4. (Existence and Uniqueness of Integral Curves) *Let $X \in \mathcal{V}(M)$ and $p \in M$. Then there exists a unique maximal integral curve $\gamma_p\colon I_p \to M$ with $\gamma_p(0) = p$.*

If $q = \gamma_p(t)$ is a point on the unique maximal integral curve of X through $p \in M$, then $I_q = I_p - t$ and

$$\gamma_q(s) := \gamma_p(t + s)$$

is the unique maximal integral curve through q. Here I_p denotes the domain of definition of the maximal integral curve through p and I_q is the domain of the maximal integral curve through q.

Example 5.3.5. (a) On $M = \mathbb{R}$ we consider the vector field X given by the function $F(s) = 1 + s^2$, i.e., $X(s) = (s, 1 + s^2)$. The corresponding ODE is

$$\gamma'(s) = X(\gamma(s)) = 1 + \gamma(s)^2.$$

For $\gamma(0) = 0$, the function $\gamma(s) := \tan(s)$ on $I :=]-\frac{\pi}{2}, \frac{\pi}{2}[$ is the unique maximal solution because

$$\lim_{t \to \frac{\pi}{2}} \tan(t) = \infty \quad \text{and} \quad \lim_{t \to -\frac{\pi}{2}} \tan(t) = -\infty.$$

(b) Let $M :=]-1, 1[$ and $X(s) = (s, 1)$, so that the corresponding ODE is $\gamma'(s) = 1$. Then the unique maximal solution is

$$\gamma(s) = s, \quad I =]-1, 1[.$$

For $M = \mathbb{R}$ the same vector field has the maximal integral curve

$$\gamma(s) = s, \quad I = \mathbb{R}.$$

The preceding example shows in particular that the global existence of integral curves can be destroyed by deleting parts of the manifold M, i.e., by considering $M' := M \setminus K$ for some closed subset $K \subseteq M$.

(c) (Linear vector fields) Let E be a Banach space. Then a vector field $X \in \mathcal{V}(E)$ is said to be *linear* if it is represented by a linear function, i.e., $X(v) = (v, Av)$ for some $A \in \mathcal{L}(E)$. The corresponding ODE is

$$\gamma'(t) = A\gamma(t).$$

We know already that the unique solutions of this ODE with the initial value $p \in E$ is of the form

$$\gamma(t) = e^{tA}p.$$

Definition 5.3.6. A vector field $X \in \mathcal{V}(M)$ is said to be *complete* if all its maximal integral curves are defined on all of $\mathbb{R}$. We write $\mathcal{V}(M)_c \subseteq \mathcal{V}(M)$ for the subset of complete vector fields.

Definition 5.3.7. Let M be a smooth manifold. A *flow on M* is a smooth map

$$\Phi\colon \mathbb{R} \times M \to M, \quad (t, m) \mapsto \Phi_t(m)$$

such that

$$\Phi_0 = \mathrm{id}_M \quad \text{and} \quad \Phi_t \circ \Phi_s = \Phi_{t+s} \quad \text{for} \quad t, s \in \mathbb{R}.$$

Writing $\mathrm{Diff}(M)$ for the group of diffeomorphisms of M, the latter conditions mean that

$$\widehat{\Phi}\colon \mathbb{R} \to \mathrm{Diff}(M), \quad t \mapsto \Phi_t$$

is a group homomorphism, i.e., a one-parameter group. A flow is the same as a *smooth action of the Lie group $\mathbb{R}$ on M* (cf. Definition 6.1.7).

Lemma 5.3.8. *If Φ is a flow, then*

$$X^\Phi(m) := \frac{d}{dt}\Big|_{t=0} \Phi(t, m) = T_{(t,m)}\Phi(1, 0)$$

defines a smooth vector field.

It is called the *velocity field* of the flow Φ.

Lemma 5.3.9. *If Φ is a flow on M, then the curves $\gamma_m(t) := \Phi(t, m)$ are integral curves of the vector field X^Φ. In particular, the flow Φ is uniquely determined by the vector field X^Φ.*

Proof. For $s, t \in \mathbb{R}$ we have

$$\gamma_x(s + t) = \Phi(s + t, x) = \Phi\big(t, \Phi(s, x)\big) = \Phi\big(t, \gamma_x(s)\big),$$

so that taking derivatives in $t = 0$ leads to $\gamma_x'(s) = X^\Phi(\gamma_x(s))$.

That Φ is uniquely determined by the vector field X^Φ follows from the uniqueness of integral curves (Theorem 5.3.4). □

For the proof of the following theorem we also refer to [La99, Thm. 4.2.6]. It is a special case of a more general result on the existence of "flows" for any smooth vector field.

Theorem 5.3.10. *Each complete vector field X is the velocity field of a unique flow Φ^X on M.*

Since the flow Φ^X is determined by the vector field X, we also call X the *infinitesimal generator of* Φ. In this sense the smooth $\mathbb{R}$-actions on a manifold M (=flows) are in one-to-one correspondence with the complete vector fields on M.

Proposition 5.3.11. (Smooth Dependence Theorem) *Let M and N be smooth manifolds and $\Psi\colon N \to \mathcal{V}(M)_c$ be a map for which the map*

$$\widehat{\Psi}\colon N \times M \to T(M), \quad (n,p) \mapsto \Psi(n)(p)$$

is smooth (the vector field $\Psi(n)$ depends smoothly on the parameter n). Then the map

$$N \times \mathbb{R} \times M \to M, \quad (n,t,m) \mapsto \Phi^{\Psi(n)}(t,m)$$

is smooth.

Proof. The parameters do not cause any additional problems, as can be seen by the following trick: On the product manifold $N \times M$ we consider the smooth vector field Y, given by

$$Y(n,p) := (0_n, \Psi(n)(p)) \in T_n(N) \times T_p(M) \cong T_{(n,p)}(N \times M).$$

Then the integral curves of Y are of the form

$$\gamma(t) = (n, \gamma_n(t)),$$

where γ_n is an integral curve of the smooth vector field $\Psi(n)$ on M. Therefore the assertion is an immediate consequence on the smoothness of the flow of Y on $N \times M$ (Theorem 5.3.10). □

Lie Derivatives

We take a closer look at the interaction of flows and vector fields. Let $X \in \mathcal{V}(M)_c$ and $Y \in \mathcal{V}(M)$. Then we define the *Lie derivative of Y along the flow of X* by

$$\mathcal{L}_X Y := \lim_{t \to 0} \frac{1}{t}((\Phi^X_{-t})_* Y - Y) = \frac{d}{dt}\Big|_{t=0} (\Phi^X_{-t})_* Y.$$

Theorem 5.3.12. $\mathcal{L}_X Y = [X,Y]$ *for $X, Y \in \mathcal{V}(M)$.*

Proof. It suffices to show that $\mathcal{L}_X Y$ and $[X,Y]$ coincide in each $p \in M$. We may therefore work in some E-chart of M.

Identifying vector fields with smooth E-valued functions, we have

$$[X,Y](x) = \mathtt{d}Y(x)X(x) - \mathtt{d}X(x)Y(x), \quad x \in U$$

(Lemma 5.2.5). On the other hand,

$$\begin{aligned} ((\Phi^X_{-t})_* Y)(x) &= T(\Phi^X_{-t}) \circ Y \circ \Phi^X_t(x) \\ &= \mathtt{d}(\Phi^X_{-t})(\Phi^X_t(x)) Y(\Phi^X_t(x)) = \big(\mathtt{d}(\Phi^X_t)(x)\big)^{-1} Y(\Phi^X_t(x)). \end{aligned}$$

To calculate the derivative of this expression with respect to t, we first observe that it does not matter if we first take derivatives with respect to t and then with respect to x or vice versa. This leads to

$$\frac{d}{dt}\Big|_{t=0} \mathtt{d}(\Phi^X_t)(x) = \mathtt{d}\Big(\frac{d}{dt}\Big|_{t=0} \Phi^X_t\Big)(x) = \mathtt{d}X(x).$$

Next we note that, for any smooth curve $\alpha\colon [-\varepsilon, \varepsilon] \to \mathrm{GL}(E)$ with $\alpha(0) = \mathbf{1}$, we have

$$(\alpha^{-1})'(t) = -\alpha(t)^{-1}\alpha'(t)\alpha(t)^{-1},$$

and in particular $(\alpha^{-1})'(0) = -\alpha'(0)$ (cf. Example 4.1.12). Combining all this, we obtain with the Product Rule

$$\mathcal{L}_X(Y)(x) = -\mathtt{d}X(x)Y(x) + \mathtt{d}Y(x)X(x) = [X,Y](x). \qquad \square$$

Corollary 5.3.13. *If $X, Y \in \mathcal{V}(M)$ are complete vector fields, then their flows $\Phi^X, \Phi^Y : \mathbb{R} \to \mathrm{Diff}(M)$ commute, i.e.,*

$$\Phi^X(t) \circ \Phi^Y(s) = \Phi^Y(s) \circ \Phi^X(t) \quad \textit{for } t, s \in \mathbb{R},$$

if and only if X and Y commute, i.e., $[X,Y] = 0$.

Proof. (1) Suppose first that Φ^X and Φ^Y commute. Let $p \in M$ and $\gamma_p(s) := \Phi^Y_s(p)$ be the integral curve of Y through p. We then have

$$\gamma_p(s) = \Phi^Y_s(p) = \Phi^X_t \circ \Phi^Y_s \circ \Phi^X_{-t}(p),$$

and passing to the derivative in $s = 0$ yields

$$Y(p) = \gamma_p'(0) = T(\Phi^X_t)Y(\Phi^X_{-t}(p)) = \big((\Phi^X_t)_*Y\big)(p).$$

Passing now to the derivative in $t = 0$, we arrive with Theorem 5.3.12 at $[X,Y] = \mathcal{L}_X(Y) = 0$.

(2) Now we assume $[X,Y] = 0$. First we show that $(\Phi^X_t)_*Y = Y$ holds for all $t \in \mathbb{R}$. For $t, s \in \mathbb{R}$ we have

$$(\Phi^X_{t+s})_*Y = (\Phi^X_t)_*(\Phi^X_s)_*Y,$$

so that

$$\frac{d}{dt}(\Phi^X_t)_*Y = -(\Phi^X_t)_*\mathcal{L}_X(Y) = 0$$

for each $t \in \mathbb{R}$. Since, for each $p \in M$, the curve

$$\mathbb{R} \to T_p(M), \quad t \mapsto \big((\Phi^X_t)_*Y\big)(p)$$

is smooth, and its derivative vanishes, it is constant equal to $Y(p)$. This shows that $(\Phi^X_t)_*Y = Y$ for each $t \in \mathbb{R}$.

For $\gamma(s) := \Phi^X_t\Phi^Y_s(p)$ we now have $\gamma(0) = \Phi^X_t(p)$ and

$$\gamma'(s) = T(\Phi^X_t) \circ Y(\Phi^Y_s(p)) = Y(\Phi^X_t\Phi^Y_s(p)) = Y(\gamma(s)),$$

so that γ is an integral curve of Y. We conclude that $\gamma(s) = \Phi^Y_s(\Phi^X_t(p))$, and this means that the flows of X and Y commute. $\square$

Remark 5.3.14. Let $X, Y \in \mathcal{V}(M)$ be two complete vector fields and Φ^X, resp., Φ^Y their flows. We then consider the commutator map

$$F\colon \mathbb{R}^2 \to \mathrm{Diff}(M), \quad (t,s) \mapsto \Phi^X_t \circ \Phi^Y_s \circ \Phi^X_{-t} \circ \Phi^Y_{-s}.$$

We know from Corollary 5.3.13 that it vanishes if and only if $[X,Y] = 0$, but there is also a more direct way from F to the Lie bracket. We first observe that the relation $\Phi^X_t \circ \Phi^Y_s \circ \Phi^X_{-t} = \Phi^{(\Phi^X_t)_*Y}_s$ (Exercise 5.3.3) leads to

$$\frac{\partial F}{\partial s}(t,0) = (\Phi^X_t)_*Y - Y \quad \text{for} \quad t \in \mathbb{R},$$ [4]

and hence with Theorem 5.3.12 to

$$\frac{\partial^2 F}{\partial t \partial s}(0,0) = [Y,X].$$

Exercises for Section 5.3

Exercise 5.3.1. Let $\varphi\colon M \to N$ be a smooth map and $X \in \mathcal{V}(M)$, $Y \in \mathcal{V}(N)$ be φ-related vector fields. Show that for any integral curve $\gamma\colon I \to M$ of X, the curve $\varphi \circ \gamma\colon I \to N$ is an integral curve of Y.

Exercise 5.3.2. Let $X \in \mathcal{V}(M)$ be a vector field and write $X^{\mathbb{R}} \in \mathcal{V}(\mathbb{R})$ for the vector field on $\mathbb{R}$, given by $X^{\mathbb{R}}(t) = (t,1)$. Show that, for an open interval $I \subseteq \mathbb{R}$, a smooth curve $\gamma\colon I \to M$ is an integral curve of X if and only if $X^{\mathbb{R}}$ and X are γ-related.

Exercise 5.3.3. Let $X \in \mathcal{V}(M)_c$ be a complete vector field and $\varphi \in \mathrm{Diff}(M)$. Then φ_*X is also complete and

$$\Phi^{\varphi_*X}_t = \varphi \circ \Phi^X_t \circ \varphi^{-1} \quad \text{for} \quad t \in \mathbb{R}.$$

Exercise 5.3.4. Let M be a smooth manifold, $\varphi \in \mathrm{Diff}(M)$ and $X \in \mathcal{V}(M)_c$ be a complete vector field. Show that the following are equivalent:

(1) φ commutes with the flow maps Φ^X_t.

(2) For each integral curve $\gamma\colon I \to M$ of X, the curve $\varphi \circ \gamma$ also is an integral curve of X.

(3) $X = \varphi_*X = T(\varphi) \circ X \circ \varphi^{-1}$, i.e., X is φ-invariant.

Exercise 5.3.5. Let $X, Y \in \mathcal{V}(M)$ be two commuting complete vector fields, i.e., $[X,Y] = 0$. Show that the vector field $X + Y$ is complete and that its flow is given by

$$\Phi^{X+Y}_t = \Phi^X_t \circ \Phi^Y_t \quad \text{for all} \quad t \in \mathbb{R}.$$

[4] Here we use that if $I \subseteq \mathbb{R}$ is an interval and $\alpha\colon I \to \mathrm{Diff}(M), \beta\colon I \to \mathrm{Diff}(M)$ are maps for which

$$\widehat{\alpha}\colon I \times M \to M, \quad (t,x) \mapsto \alpha(t)(x) \quad \text{and} \quad \widehat{\beta}\colon I \times M \to M, \quad (t,x) \mapsto \beta(t)(x)$$

are smooth, then the curve $\gamma(t) := \alpha(t) \circ \beta(t)$ also has this property (by the Chain Rule), and if $\alpha(0) = \beta(0) = \mathrm{id}_M$, then γ satisfies

$$\gamma'(0) = \alpha'(0) \circ \beta(0) + T(\alpha(0)) \circ \beta'(0) = \alpha'(0) + \beta'(0).$$

Chapter 6

Lie Algebra and Exponential Function of a Lie Group

In this chapter we introduce the Lie algebra and the exponential function of a general Lie group (modeled on a Banach space). The Lie algebra is obtained from the space of left invariant vector fields. Since all these vector fields are complete, one obtains the exponential function from their flows. The Lie functor $\mathbf{L}\colon G \mapsto \mathbf{L}(G)$ assigns a Banach–Lie algebra to each Lie group and a Lie algebra homomorphism $\mathbf{L}(\varphi)$ to each morphism φ of Lie groups. It is the key tool to translate Lie group problems into problems in linear algebra.

6.1 The Lie Algebra of a Lie group

Before we turn to the definition of the Lie algebra of a Lie group, we show that the group structure on a Lie group G induced a natural Lie group structure on its tangent bundle TG.

In the following we always identify the tangent bundle $T(G \times G)$ with the product space $TG \times TG$ (Remark 5.1.7).

Lemma 6.1.1. (a) *The tangent map*

$$T(m_G)\colon T(G \times G) \cong T(G) \times T(G) \to T(G), \quad (v, w) \mapsto v \cdot w := Tm_G(v, w)$$

defines a Lie group structure on $T(G)$ with identity element $0_{\mathbf{1}} \in T_{\mathbf{1}}(G)$ and inversion $T(\eta_G)$. For $v \in T_g(G)$ and $w \in T_h(G)$, we have

$$v \cdot w = T_g(\rho_h)v + T_h(\lambda_g)w = v \cdot 0_h + 0_g \cdot w. \tag{6.1}$$

(b) *The canonical projection $\pi_{T(G)}\colon T(G) \to G$ is a morphism of Lie groups with kernel $(T_{\mathbf{1}}(G), +)$ and the zero section $\sigma\colon G \to T(G), g \mapsto 0_g \in T_g(G)$ is a homomorphism of Lie groups with $\pi_{T(G)} \circ \sigma = \mathrm{id}_G$.*

(c) *The map*

$$\Phi\colon G \times T_{\mathbf{1}}(G) \to T(G), \quad (g, x) \mapsto g.x := 0_g \cdot x = T(\lambda_g)x$$

is a diffeomorphism.

Proof. (a) Since the multiplication map $m_G\colon G \times G \to G$ is smooth, the same holds for its tangent map

$$Tm_G\colon T(G \times G) \cong T(G) \times T(G) \to T(G).$$

Let $\varepsilon_G\colon G \to G, g \mapsto \mathbf{1}$ be the constant homomorphism. Then the group axioms for G are encoded in the relations

(1) $m_G \circ (m_G \times \mathrm{id}_G) = m_G \circ (\mathrm{id}_G \times m_G)$ (associativity),

(2) $m_G \circ (\eta_G, \mathrm{id}_G) = m_G \circ (\mathrm{id}_G, \eta_G) = \varepsilon_G$ (inversion), and

(3) $m_G \circ (\varepsilon_G, \mathrm{id}_G) = m_G \circ (\mathrm{id}_G, \varepsilon_G) = \mathrm{id}_G$ (unit element).

Using the functoriality of T and its compatibility with products (cf. Remark 5.1.8), we see that these properties carry over to the corresponding maps on $T(G)$:

(1) $T(m_G) \circ T(m_G \times \mathrm{id}_G) = T(m_G) \circ (T(m_G) \times \mathrm{id}_{T(G)})$
$= T(m_G) \circ (\mathrm{id}_{T(G)} \times T(m_G))$ (associativity),

(2) $T(m_G) \circ (T(\eta_G), \mathrm{id}_{T(G)}) = T(m_G) \circ (\mathrm{id}_{T(G)}, T(\eta_G)) = T(\varepsilon_G)$ (inversion), and

(3) $T(m_G) \circ (T(\varepsilon_G), \mathrm{id}_{T(G)}) = T(m_G) \circ (\mathrm{id}_{T(G)}, T(\varepsilon_G)) = \mathrm{id}_{T(G)}$ (unit element).

Here we only have to observe that the tangent map $T(\varepsilon_G)$ maps each $v \in T(G)$ to $0_{\mathbf{1}} \in T_{\mathbf{1}}(G)$, which is the neutral element of $T(G)$. We conclude that $T(G)$ is a Lie group with multiplication $T(m_G)$, inversion $T(\eta_G)$, and unit element $0_{\mathbf{1}} \in T_{\mathbf{1}}(G)$.

For $v \in T_g(G)$ and $w \in T_h(G)$, the linearity of $T_{(g,h)}(m_G)$ implies that

$$\begin{aligned} Tm_G(v, w) &= T_{(g,h)}(m_G)(v, w) = T_{(g,h)}(m_G)(v, 0) + T_{(g,h)}(m_G)(0, w) \\ &= T_g(\rho_h)v + T_h(\lambda_g)w, \end{aligned}$$

(b) The definition of the tangent map implies that the zero section $\sigma\colon G \to T(G)$ satisfies

$$Tm_G \circ (\sigma \times \sigma) = \sigma \circ m_G \quad \text{and} \quad Tm_G(0_g, 0_h) = 0_{m_G(g,h)} = 0_{gh},$$

which means that it is a morphism of Lie groups. That $\pi_{T(G)}$ also is a morphism of Lie groups follows likewise from the relation

$$\pi_{T(G)} \circ Tm_G = m_G \circ (\pi_{T(G)} \times \pi_{T(G)}),$$

which also is an immediate consequence of the definition of the tangent map T_{m_G}: it maps $T_g(G) \times T_h(G)$ into $T_{gh}(G)$ (cf. Exercise 5.1.4). From (6.1), we obtain in particular that the multiplication on the normal subgroup $\ker \pi_{T(G)} = T_{\mathbf{1}}(G)$ is simply given by addition.

(c) The smoothness of Φ follows from the smoothness of the multiplication of $T(G)$ and the smoothness of the zero section $\sigma\colon G \to T(G), g \mapsto 0_g$. That Φ is a diffeomorphism follows from the following explicit formula for its inverse: $\Phi^{-1}(v) = (\pi_{T(G)}(v), \pi_{T(G)}(v)^{-1}v)$, so that its smoothness follows from the smoothness of $\pi_{T(G)}$ (its first component), and the smoothness of the multiplication on $T(G)$. □

Definition 6.1.2. In the following we shall mostly use the simplified notation

$$g.v := 0_g \cdot v \quad \text{for} \quad g \in G, v \in TG.$$

We likewise write

$$v.g := v \cdot 0_g \quad \text{for} \quad g \in G, v \in TG.$$

Definition 6.1.3. (The Lie algebra of G) A vector field $X \in \mathcal{V}(G)$ is called *left invariant* if

$$X = (\lambda_g)_* X := T(\lambda_g) \circ X \circ \lambda_g^{-1} \quad \text{for} \quad g \in G.$$

We write $\mathcal{V}(G)^l$ for the set of left invariant vector fields in $\mathcal{V}(G)$. Clearly $\mathcal{V}(G)^l$ is a linear subspace of $\mathcal{V}(G)$. The left invariance of X means that X is λ_g-related to itself for every $g \in G$. Therefore the Related Vector Field Lemma 5.2.8 implies that, if X and Y are left-invariant, then their Lie bracket $[X, Y]$ is also λ_g-related to itself for each $g \in G$, hence left invariant. We conclude that the vector space $\mathcal{V}(G)^l$ is a Lie subalgebra of $(\mathcal{V}(G), [\cdot,\cdot])$.

Next we observe that the left invariance of a vector field X implies that for each $g \in G$ we have $X(g) = g.X(\mathbf{1})$ (Lemma 6.1.1(b)), so that X is completely determined by its value $X(\mathbf{1}) \in T_{\mathbf{1}}(G)$. Conversely, for each $x \in T_{\mathbf{1}}(G)$, we obtain a left invariant vector field $x_l \in \mathcal{V}(G)^l$ with $x_l(\mathbf{1}) = x$ by $x_l(g) := g.x$. That this vector field is indeed left invariant follows from

$$x_l \circ \lambda_h(g) = x_l(hg) = (hg).x = h.(g.x) = T(\lambda_h)x_l(g)$$

for all $h, g \in G$. Hence

$$T_{\mathbf{1}}(G) \to \mathcal{V}(G)^l, \quad x \mapsto x_l$$

is a linear bijection. We thus obtain a Lie bracket $[\cdot,\cdot]$ on $T_{\mathbf{1}}(G)$ satisfying

$$[x, y] = [x_l, y_l](\mathbf{1}) \quad \text{and} \quad [x, y]_l = [x_l, y_l] \quad \text{for all} \quad x, y \in T_{\mathbf{1}}(G). \tag{6.2}$$

To show that the Lie bracket on the Banach space $T_{\mathbf{1}}(G)$ is continuous, let $E := T_{\mathbf{1}}(G)$ and choose a local E-chart (φ, U) of G with $\varphi(\mathbf{1}) = 0$ and $T_{\mathbf{1}}(\varphi) = \mathrm{id}_E$. For $x \in T_{\mathbf{1}}(G)$ we then obtain a smooth vector field

$$\widetilde{x}_l := \varphi_* x_l = T(\varphi) \circ x_l \circ \varphi^{-1} \tag{6.3}$$

on $V := \varphi(U)$. We identify $T(V) \cong V \times E$ and vector fields on V with smooth E-valued functions. Then the Related Vector Field Lemma 5.2.8 implies

$$\begin{aligned} [x, y] &= [x_l, y_l](\mathbf{1}) = [\widetilde{x}_l, \widetilde{y}_l](0) = \mathtt{d}\widetilde{y}_l(0)\widetilde{x}_l(0) - \mathtt{d}\widetilde{x}_l(0)\widetilde{y}_l(0) \\ &= \mathtt{d}\widetilde{y}_l(0)x - \mathtt{d}\widetilde{x}_l(0)y. \end{aligned} \tag{6.4}$$

Clearly, the function

$$\psi\colon E \times V \to E, \quad (x,z) \mapsto \widetilde{x}_l(z) = T(\varphi)x_l(\varphi^{-1}(z)) = T(\varphi)T(m_G)(0_{\varphi^{-1}(z)}, x)$$

is smooth, so that $(x,y) \mapsto \mathtt{d}\widetilde{x}_l(0)y$ is continuous bilinear, and hence the bracket on $E \cong T_{\mathbf{1}}(G)$ is continuous.

The Banach–Lie algebra

$$\mathbf{L}(G) := (T_{\mathbf{1}}(G), [\cdot,\cdot]) \cong \mathcal{V}(G)^l$$

is called *the Lie algebra of G*.

Proposition 6.1.4. (Functoriality of the Lie algebra) *If $\varphi\colon G \to H$ is a morphism of Lie groups, then the tangent map*

$$\mathbf{L}(\varphi) := T_{\mathbf{1}}(\varphi)\colon \mathbf{L}(G) \to \mathbf{L}(H)$$

is a continuous homomorphism of Banach–Lie algebras.

Proof. Let $x, y \in \mathbf{L}(G)$ and x_l, y_l be the corresponding left invariant vector fields. Then $\varphi \circ \lambda_g = \lambda_{\varphi(g)} \circ \varphi$ for each $g \in G$ implies that

$$T(\varphi) \circ T(\lambda_g) = T(\lambda_{\varphi(g)}) \circ T(\varphi),$$

and applying this relation to $x, y \in T_{\mathbf{1}}(G)$, we get

$$T\varphi \circ x_l = \big(\mathbf{L}(\varphi)x\big)_l \circ \varphi \quad \text{and} \quad T\varphi \circ y_l = \big(\mathbf{L}(\varphi)y\big)_l \circ \varphi, \tag{6.5}$$

i.e., x_l is φ-related to $\big(\mathbf{L}(\varphi)x\big)_l$ and y_l is φ-related to $\big(\mathbf{L}(\varphi)y\big)_l$. Therefore the Related Vector Field Lemma 5.2.8 implies that

$$T\varphi \circ [x_l, y_l] = [\big(\mathbf{L}(\varphi)x\big)_l, \big(\mathbf{L}(\varphi)y\big)_l] \circ \varphi.$$

Evaluating at $\mathbf{1}$, we obtain $\mathbf{L}(\varphi)[x,y] = [\mathbf{L}(\varphi)(x), \mathbf{L}(\varphi)(y)]$, showing that $\mathbf{L}(\varphi)$ is a homomorphism of Lie algebras. That it is continuous follows from the smoothness of $T\varphi$. □

Remark 6.1.5. We obviously have $\mathbf{L}(\mathrm{id}_G) = \mathrm{id}_{\mathbf{L}(G)}$, and for two morphisms $\varphi_1\colon G_1 \to G_2$ and $\varphi_2\colon G_2 \to G_3$ of Lie groups, we obtain

$$\mathbf{L}(\varphi_2 \circ \varphi_1) = \mathbf{L}(\varphi_2) \circ \mathbf{L}(\varphi_1),$$

from the Chain Rule:

$$T_{\mathbf{1}}(\varphi_2 \circ \varphi_1) = T_{\varphi_1(\mathbf{1})}(\varphi_2) \circ T_{\mathbf{1}}(\varphi_1) = T_{\mathbf{1}}(\varphi_2) \circ T_{\mathbf{1}}(\varphi_1).$$

The preceding lemma implies that the assignments $G \mapsto \mathbf{L}(G)$ and $\varphi \mapsto \mathbf{L}(\varphi)$ define a functor, called the *Lie functor*,

$$\mathbf{L}\colon \underline{LieGrp} \to \underline{LieAlg}$$

from the category $\underline{LieGrp}$ of Lie groups to the category $\underline{BLieAlg}$ of Banach–Lie algebras.

Corollary 6.1.6. *For each isomorphism of Lie groups $\varphi\colon G \to H$, the map $\mathbf{L}(\varphi)$ is an isomorphism of Banach–Lie algebras, and for each $x \in \mathbf{L}(G)$, the following equation holds*

$$\varphi_* x_l := T(\varphi) \circ x_l \circ \varphi^{-1} = \big(\mathbf{L}(\varphi)x\big)_l. \tag{6.6}$$

Proof. Let $\psi\colon H \to G$ be the inverse of φ. Then $\varphi \circ \psi = \mathrm{id}_H$ and $\psi \circ \varphi = \mathrm{id}_G$ leads to $\mathbf{L}(\varphi) \circ \mathbf{L}(\psi) = \mathrm{id}_{\mathbf{L}(H)}$ and $\mathbf{L}(\psi) \circ \mathbf{L}(\varphi) = \mathrm{id}_{\mathbf{L}(G)}$ (Remark 6.1.5). Further (6.6) follows from (6.5) in the proof of Proposition 6.1.4. □

Smooth Actions of Lie Groups

We already encountered smooth flows on manifolds in Section 5.3. These can be viewed as actions of the one-dimensional Lie group $(\mathbb{R}, +)$. In particular, we have seen that these actions are in one-to-one correspondence with complete vector fields, which is the corresponding Lie algebra picture. Now we describe the corresponding concept for general Lie groups.

Definition 6.1.7. 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 $g_1, g_2 \in G$ and $m \in M$.

We also write

$$g.m := \sigma(g, m), \quad \sigma_g(m) := \sigma(g, m) \quad \text{and} \quad \sigma^m(g) := \sigma(g, m) = g.m.$$

The map σ^m is called the *orbit map*.

For each smooth action σ, the map

$$\widehat{\sigma}\colon G \to \mathrm{Diff}(M), \quad g \mapsto \sigma_g$$

is a group homomorphism. Conversely, 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.

Remark 6.1.8. What we call an action is sometimes called a *left action*. Likewise one defines a *right action* as a smooth map $\sigma_R\colon M \times G \to M$ 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 $m.g := \sigma_R(m, g)$, 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 left action. This translation is one-to-one, so that we may freely pass from one type of action to the other.

Examples 6.1.9. (a) If $X \in \mathcal{V}(M)$ is a complete vector field (cf. Definition 5.3.6) and $\Phi\colon \mathbb{R} \times M \to M$ its flow, then Φ defines a smooth action of $G = (\mathbb{R}, +)$ on M.
(b) If G is a Lie group, then the multiplication map $\sigma := m_G\colon G \times G \to G$ defines a smooth left action of G on itself. In this case the $(m_G)_g = \lambda_g$ are the left multiplications.

The multiplication map also defines a smooth right action of G on itself. The corresponding left action is

$$\sigma\colon G \times G \to G, \quad (g, h) \mapsto hg^{-1} \quad \text{with} \quad \sigma_g = \rho_g^{-1}.$$

There is a third action of G on itself, the *conjugation action*:

$$\sigma\colon G \times G \to G, \quad (g, h) \mapsto ghg^{-1} \quad \text{with} \quad \sigma_g = c_g.$$

(c) For every Banach space E, we have a natural smooth action of the Lie group $\mathrm{GL}(E)$ on E:

$$\sigma\colon \mathrm{GL}(E) \times E \to E, \quad \sigma(g, x) := gx.$$

We further have a smooth action of $\mathrm{GL}(E)$ on $\mathcal{L}(E)$:

$$\sigma\colon \mathrm{GL}(E) \times \mathcal{L}(E) \to \mathcal{L}(E), \quad \sigma(g, A) = gAg^{-1}.$$

Note that this example specializes to $E = \mathbb{R}^n$, where we obtain actions of $\mathrm{GL}_n(\mathbb{R})$ on $\mathbb{R}^n$ and $M_n(\mathbb{R}) \cong \mathcal{L}(\mathbb{R}^n)$.
(d) For two Banach spaces E and F, we obtain on the Banach space $\mathcal{L}(E, F)$ a smooth action of the product Lie group $\mathrm{GL}(F) \times \mathrm{GL}(E)$ by $\sigma((g, h), A) := gAh^{-1}$.

For $E = \mathbb{R}^q$ and $F = \mathbb{R}^p$, the space $\mathcal{L}(E, F)$ can be identified with the space $M_{p,q}(\mathbb{R})$ of $(p\times q)$-matrices, on which the Lie group $\mathrm{GL}_p(\mathbb{R})\times\mathrm{GL}_q(\mathbb{R})$ acts by $\sigma((g, h), A) := gAh^{-1}$.

The following proposition generalizes the passage from flows of vector fields to actions of general Lie groups. Specializing it to a smmoth flow $\sigma = \Phi\colon \mathbb{R} \times M \to M$, the vector field $-\dot\sigma(1) = X^\Phi$ is the infinitesimal generator of the flow.

Proposition 6.1.10. *Let G be a Lie group and $\sigma\colon G \times M \to M$ be a smooth action of G on M. Then the assignment*

$$\dot\sigma\colon \mathbf{L}(G) \to \mathcal{V}(M), \qquad \dot\sigma(x)(m) := \mathbf{L}(\sigma)(x)(m) := -T_{\mathbf{1}}(\sigma^m)(x)$$

is a homomorphism of Lie algebras.

Proof. First we observe that, for each $x \in \mathbf{L}(G)$, the map $\dot\sigma(x)$ defines a smooth map $M \to T(M)$, and since $\dot\sigma(x)(m) \in T_{\sigma(\mathbf{1},m)}(M) = T_m(M)$, it is a smooth vector field on M.

To see that $\dot\sigma$ is a homomorphism of Lie algebras, we pick $m \in M$ and write

$$\varphi^m := \sigma^m \circ \eta_G \colon G \to M, \quad g \mapsto g^{-1}.m$$

for the reversed orbit map. Then

$$\varphi^m(gh) = (gh)^{-1}.m = h^{-1}.(g^{-1}.m) = \varphi^{g^{-1}.m}(h),$$

which can be written as

$$\varphi^m \circ \lambda_g = \varphi^{g^{-1}.m}.$$

Taking the differential in $\mathbf{1} \in G$, we obtain for each $x \in \mathbf{L}(G) = T_\mathbf{1}(G)$:

$$\begin{aligned} T_g(\varphi^m)x_l(g) &= T_g(\varphi^m)T_\mathbf{1}(\lambda_g)x = T_\mathbf{1}(\varphi^m \circ \lambda_g)x = T_\mathbf{1}(\varphi^{g^{-1}.m})x \\ &= T_\mathbf{1}(\sigma^{g^{-1}.m})T_\mathbf{1}(\eta_G)x = -T_\mathbf{1}(\sigma^{\varphi^m(g)})x = \dot\sigma(x)(\varphi^m(g)). \end{aligned}$$

This means that the left invariant vector field x_l on G is φ^m-related to the vector field $\dot\sigma(x)$ on M. Therefore the Related Vector Field Lemma 5.2.8 implies that for $x, y \in \mathbf{L}(G)$ the vector field $[x_l, y_l]$ is φ^m-related to $[\dot\sigma(x), \dot\sigma(y)]$, which leads for each $m \in M$ to

$$\begin{aligned} \mathbf{L}(\sigma)([x,y])(m) &= T_\mathbf{1}(\varphi^m)[x,y]_l(\mathbf{1}) = T_\mathbf{1}(\varphi^m)[x_l,y_l](\mathbf{1}) \\ &= [\dot\sigma(x), \dot\sigma(y)](\varphi^m(\mathbf{1})) = [\dot\sigma(x), \dot\sigma(y)](m). \end{aligned}$$ □

Exercises for Section 6.1

Exercise 6.1.1. Show that the natural group structure on $\mathbb{T} \cong \mathbb{S}^1 \subseteq \mathbb{C}^\times$ turns it into a Lie group.

Exercise 6.1.2. Let $G_1, \ldots, G_n$ be Lie groups and $G := G_1 \times \ldots \times G_n$, endowed with the direct product group structure

$$(g_1, \ldots, g_n)(g_1', \ldots, g_n') := (g_1 g_1', \ldots, g_n g_n')$$

and the product manifold structure. Show that G is a Lie group with

$$\mathbf{L}(G) \cong \mathbf{L}(G_1) \times \ldots \times \mathbf{L}(G_n).$$

Exercise 6.1.3. Let V and W be Banach spaces and $\beta\colon V \times V \to W$ be a bilinear map. Show that $G := W \times V$ is a Lie group with respect to

$$(w, v)(w', v') := (w + w' + \beta(v, v'), v + v').$$

For $(w, v) \in \mathbf{L}(G) \cong T_{(0,0)}(G)$, find a formula for the corresponding left invariant vector field $(w, v)_l$, considered as a smooth function $G \to W \times V$.

Exercise 6.1.4. (Automatic smoothness of the inversion) Let G be a manifold, endowed with a group structure for which the multiplication map m_G is smooth. Show that:

(1) $T_{(g,h)}(m_G) = T_g(\rho_h) + T_h(\lambda_g)$ for $\lambda_g(x) = gx$ and $\rho_h(x) = xh$.

(2) $T_{(\mathbf{1},\mathbf{1})}(m_G)(v, w) = v + w$.

(3) The inverse map $\eta_G \colon G \to G, g \mapsto g^{-1}$ is smooth if it is smooth in a neighborhood of $\mathbf{1}$. Hint: Consider the smooth map

$$F \colon G \times G \to G \times G, uad(x, y) \mapsto (xy, y)$$

and apply the Inverse Function Theorem in $(\mathbf{1}, \mathbf{1})$ (cf. Exercise 5.1.3).

(4) The inverse map η_G is smooth.

Conclude that G is a Lie group.

Exercise 6.1.5. Let $\mathcal{A}$ be a unital Banach algebra. Show that, identifying vector fields on the open subset $\mathcal{A}^\times$ with smooth $\mathcal{A}$-valued functions, a vector field $X \in \mathcal{V}(\mathcal{A}^\times) \cong C^\infty(\mathcal{A}^\times, \mathcal{A})$ is left invariant if and only if there exists an element $x \in \mathcal{A}$ with $X(g) = gx$ for $g \in \mathcal{A}^\times$.

Exercise 6.1.6. Let G be a Lie group and X a vector field on G. Endowing TG with its natural Lie group structure, show that X is left invariant if and only if

$$X(gh) = g.X(h) \quad \text{for} \quad g, h \in G.$$

Exercise 6.1.7. Let $G = E$ be a Banach space. Show that a vector fields $X \in \mathcal{V}(E) \cong C^\infty(E, E)$ is left invariant if and only if it corresponds to a constant function.

Exercise 6.1.8. Consider the three-dimensional Heisenberg group

$$G = \left\{ \begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{R}) : x, y, z \in \mathbb{R} \right\}$$

Determine the space of (left) invariant vector fields in the coordinates (x, y, z).

Exercise 6.1.9. Let $f_1, f_2 \colon G \to H$ be two group homomorphisms. Show that the pointwise product

$$f_1 f_2 \colon G \to H, \quad g \mapsto f_1(g) f_2(g)$$

is a homomorphism if and only if $f_1(G)$ commutes with $f_2(G)$.

Exercise 6.1.10. Let M be a manifold and V a finite-dimensional vector space with a basis $(b_1, \ldots, b_n)$. Let $f \colon M \to \mathrm{GL}(V)$ be a map. Show that the following are equivalent:

(1) f is smooth.

(2) For each $v \in V$, the map $f_v : M \to V, m \mapsto f(m)v$ is smooth.

(3) For each i, the map $f : M \to V, m \mapsto f(m)b_i$ is smooth.

Exercise 6.1.11. Let G be a Lie group. Show that any map $\varphi\colon G \to G$ commuting with all left multiplications λ_g, $g \in G$, is a right multiplication.

Exercise 6.1.12. Let V be a Banach space and $\mu_t(v) := tv$ for $t \in \mathbb{R}^\times$. Show that:

(1) A vector field $X \in \mathcal{V}(V)$ is linear if and only if $(\mu_t)_*X = X$ holds for all $t \in \mathbb{R}^\times$.

(2) A diffeomorphism $\varphi \in \mathrm{Diff}(V)$ is linear if and only if it commutes with all the maps μ_t, $t \in \mathbb{R}^\times$.

6.2 The Exponential Function of a Lie Group

In the preceding section we have introduced the Lie functor which assigns to a Lie group G its Lie algebra $\mathbf{L}(G)$ and to a morphism φ of Lie groups its tangent morphism $\mathbf{L}(\varphi)$ of Lie algebras. In this section, we introduce a key tool of Lie theory which will allow us to also go in the opposite direction: the exponential function

$$\exp_G\colon \mathbf{L}(G) \to G.$$

It is a natural generalization of the exponential map $\exp\colon \mathcal{A} \to \mathcal{A}^\times$ of a unital Banach algebra $\mathcal{A}$. We conclude this section with a discussion of the naturality of the exponential function (Proposition 6.2.9) and the Lie group versions of the Trotter Product Formula and the Commutator Formula.

Basic Properties of the Exponential Function

Proposition 6.2.1. *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 $X \in \mathcal{V}(G)^l$ with $\gamma(0) = g$ (cf. Theorem 5.3.4).

For each $h \in G$ we have $(\lambda_h)_*X = X$, which implies that $\eta := \lambda_h \circ \gamma$ also is an integral curve of X (Exercise 5.3.1). 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 maximality of I, it now follows that $I - s \subseteq I$, and hence that $I - ns \subseteq I$ for each $n \in \mathbb{N}$, so that the interval I is unbounded from below. Applying the same argument to some $s < 0$, we see that I is also unbounded from above. Hence $I = \mathbb{R}$, which means that X is complete. $\square$

Definition 6.2.2. We now define the *exponential function*

$$\exp_G\colon \mathbf{L}(G) \to G, \quad \exp_G(x) := \gamma_x(1) = \Phi^{x_l}(1, \mathbf{1}),$$

where $\gamma_x \colon \mathbb{R} \to G$ is the unique maximal integral curve of the left invariant vector field x_l, satisfying $\gamma_x(0) = \mathbf{1}$. This means that γ_x is the unique solution of the initial value problem

$$\gamma(0) = \mathbf{1}, \quad \gamma'(t) = x_l(\gamma(t)) = \gamma(t).x \quad \text{for} \quad t \in \mathbb{R}.$$

Example 6.2.3. (a) Let $G := (E, +)$ be the additive group of the Banach space E. The left invariant vector fields on E are given by

$$x_l(w) := \frac{d}{dt}\Big|_{t=0} w + tx = x,$$

so that they are simply the constant vector fields. Hence (cf. Lemma 5.2.5)

$$[x_l, y_l](0) = \mathtt{d}y_l(x_l(0)) - \mathtt{d}x_l(y_l(0)) = \mathtt{d}y_l(x) - \mathtt{d}x_l(y) = 0.$$

Therefore $\mathbf{L}(E)$ is an abelian Lie algebra.

For each $x \in E$, the flow of x_l is given by $\Phi^{x_l}(t, v) = v + tx$, so that

$$\exp_E(x) = \Phi^{x_l}(1, 0) = x, \quad \text{i.e.,} \quad \exp_E = \mathrm{id}_E\,.$$

(b) Now let $G := \mathcal{A}^\times$ be the unit group of a unital Banach algebra $\mathcal{A}$. The left invariant vector field A_l corresponding to an element $A \in \mathcal{A}$ is given by

$$A_l(g) = T_{\mathbf{1}}(\lambda_g)A = gA$$

because $\lambda_g(h) = gh$ extends to a linear endomorphism of $\mathcal{A}$. The unique solution $\gamma_A \colon \mathbb{R} \to \mathcal{A}^\times$ of the initial value problem

$$\gamma(0) = \mathbf{1}, \quad \gamma'(t) = A_l(\gamma(t)) = \gamma(t)A$$

is the curve

$$\gamma_A(t) = e^{tA} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k.$$

It follows that $\exp_G(A) = \gamma_A(1) = e^A$ is the exponential function of $\mathcal{A}$.

The Lie algebra $\mathbf{L}(G)$ of G is determined from

$$\begin{aligned}[A, B] = [A_l, B_l](\mathbf{1}) &= \mathtt{d}B_l(\mathbf{1})A_l(\mathbf{1}) - \mathtt{d}A_l(\mathbf{1})B_l(\mathbf{1}) \\ &= \mathtt{d}B_l(\mathbf{1})A - \mathtt{d}A_l(\mathbf{1})B = AB - BA.\end{aligned}$$

Therefore the Lie bracket on $\mathbf{L}(G) = T_{\mathbf{1}}(G) \cong \mathcal{A}$ is given by the commutator bracket, i.e., $\mathbf{L}(\mathcal{A}^\times) = \mathcal{A}_L$.

We now return to general Lie groups.

Lemma 6.2.4. *Let G be a Lie group.*

(a) *For each $x \in \mathbf{L}(G)$, the curve $\gamma_x \colon \mathbb{R} \to G, \gamma_x(t) = \exp_G(tx)$ is a smooth homomorphism of Lie groups with $\gamma_x'(0) = x$.*

(b) *The flow of the left invariant vector field x_l is given by*

$$\Phi(t, g) = g\gamma_x(t) = g \exp_G(tx).$$

(c) *If* $\gamma\colon \mathbb{R} \to G$ *is a smooth homomorphism of Lie groups and* $x := \gamma'(0)$*, then* $\gamma = \gamma_x$*. In particular, the map*

$$\mathrm{Hom}(\mathbb{R}, G) \to \mathbf{L}(G), \quad \gamma \mapsto \gamma'(0)$$

is a bijection, where $\mathrm{Hom}(\mathbb{R}, G)$ *stands for 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_l, it is a smooth curve. Hence the smoothness of the multiplication in G implies that $\Phi(t, g) := g\gamma_x(t)$ defines a smooth map $\mathbb{R} \times G \to G$. In view of the left invariance of x_l, we have for each $g \in G$ and $\Phi^g(t) := \Phi(t, g)$ the relation

$$(\Phi^g)'(t) = T(\lambda_g)\gamma_x'(t) = T(\lambda_g)x_l(\gamma_x(t)) = x_l(g\gamma_x(t)) = x_l(\Phi^g(t)).$$

Therefore Φ^g is an integral curve of x_l with $\Phi^g(0) = g$, and this proves that Φ is the unique maximal flow of the complete vector field x_l.

In particular, we obtain for $t, s \in \mathbb{R}$:

$$\gamma_x(t + s) = \Phi(t + s, \mathbf{1}) = \Phi(t, \Phi(s, \mathbf{1})) = \Phi(s, \mathbf{1})\gamma_x(t) = \gamma_x(s)\gamma_x(t). \tag{6.7}$$

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

$$\Phi(t, g) := g\gamma(t)$$

defines a flow on G whose infinitesimal generator is the vector field given by

$$X(g) = \frac{d}{dt}\Big|_{t=0} \Phi(t, g) = T(\lambda_g)\gamma'(0).$$

We conclude that $X = x_l$ for $x = \gamma'(0)$, so that X is a left invariant vector field. Since γ is its unique integral curve through 0, it follows that $\gamma = \gamma_x$. In view of (a), this proves (c). □

Proposition 6.2.5. *For a Lie group* G*, the exponential function*

$$\exp_G\colon \mathbf{L}(G) \to G$$

is smooth and satisfies

$$T_0(\exp_G) = \mathrm{id}_{\mathbf{L}(G)}.$$

In particular, $\exp_G$ *is a local diffeomorphism in* 0 *in the sense that it maps some* 0*-neighborhood in* $\mathbf{L}(G)$ *diffeomorphically onto some open* $\mathbf{1}$*-neighborhood in* G*.*

Proof. The map $\Psi\colon \mathbf{L}(G) \to \mathcal{V}(G), x \mapsto x_l$ satisfies the assumptions of Proposition 5.3.11 because the map

$$\mathbf{L}(G) \times G \to T(G), \quad (x, g) \mapsto x_l(g) = g.x$$

is smooth (Lemma 6.1.1). In the terminology of Proposition 5.3.11, it now follows that the map

$$\Phi\colon \mathbb{R}\times \mathbf{L}(G)\times G\to G,\quad (t,x,g)\mapsto g\gamma_x(t)=g\exp_G(tx)$$

is smooth. Therefore $\exp_G(x)=\Phi(1,x,\mathbf{1})$ is also smooth.

Finally, we observe that

$$T_0(\exp_G)(x)=\frac{d}{dt}\Big|_{t=0}\exp_G(tx)=\gamma_x'(0)=x,$$

so that $T_0(\exp_G)=\mathrm{id}_{\mathbf{L}(G)}$, so that the remaining assertions follow from the Inverse Function Theorem (Exercise 5.1.3). □

Lemma 6.2.6. *If $\sigma\colon G\times M\to M,(g,m)\mapsto g.m$ is a smooth action and $x\in\mathbf{L}(G)$, then the flow of the vector field $\dot\sigma(x)$ is given by $\Phi^x(t,m)=\exp_G(-tx).m$. In particular,*

$$\dot\sigma(x)(m)=\frac{d}{dt}\Big|_{t=0}\exp_G(-tx).m.$$

Proof. In the proof of Proposition 6.1.10, we have seen that

$$T_g(\varphi^m)x_l(g)=\dot\sigma(x)(\varphi^m(g))$$

holds for the map $\varphi^m(g)=g^{-1}.m$. In view of Proposition 6.2.5, this leads to

$$\frac{d}{dt}\Big|_{t=0}\exp_G(-tx).m=T_{\mathbf{1}}(\varphi^m)T_0(\exp_G)x=T_{\mathbf{1}}(\varphi^m)x=\dot\sigma(x)(m),$$

and hence proves the lemma. □

Corollary 6.2.7. *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 and y commute, then the corresponding left invariant vector fields commute, and Corollary 5.3.13 implies that their flows commute. We conclude that, for all $t,s\in\mathbb{R}$, we have

$$\exp_G(tx)\exp_G(sy)=\exp_G(sy)\exp_G(tx). \tag{6.8}$$

Therefore

$$\gamma(t):=\exp_G(tx)\exp_G(ty)$$

is a smooth group homomorphism. In view of

$$\gamma'(0)=T_{(\mathbf{1},\mathbf{1})}(m_G)(x,y)=x+y$$

(Lemma 6.1.1), Lemma 6.2.4(c) leads to $\gamma(t)=\exp_G(t(x+y))$, and for $t=1$ we obtain the lemma. □

Lemma 6.2.8. *The subgroup $\langle\exp_G(\mathbf{L}(G))\rangle$ of G generated by $\exp_G(\mathbf{L}(G))$ coincides with the identity component G_0 of G, i.e., the connected component containing $\mathbf{1}$.*

Proof. Since $\exp_G$ is a local diffeomorphism in 0 (Proposition 6.2.5), $\exp_G(\mathbf{L}(G))$ is a neighborhood of $\mathbf{1}$. We conclude that the subgroup $H := \langle \exp_G(\mathbf{L}(G)) \rangle$ generated by the exponential image is a $\mathbf{1}$-neighborhood, hence contains G_0 (cf. Lemma 1.2.2(iv)). On the other hand, $\exp_G$ is continuous, so that it maps the connected space $\mathbf{L}(G)$ into the identity component G_0 of G, which leads to $H \subseteq G_0$, and hence to equality. $\square$

Naturality of the Exponential Function

In this subsection we study how the exponential function is related to the Lie functor.

Proposition 6.2.9. *Let $\varphi\colon G_1 \to G_2$ be a morphism of Lie groups and $\mathbf{L}(\varphi)\colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ its differential in $\mathbf{1}$. Then*

$$\exp_{G_2} \circ \mathbf{L}(\varphi) = \varphi \circ \exp_{G_1}, \tag{6.9}$$

i.e., the following diagram commutes

$$\begin{array}{ccc} G_1 & \xrightarrow{\quad\varphi\quad} & G_2 \\ \uparrow{\scriptstyle \exp_{G_1}} & & \uparrow{\scriptstyle \exp_{G_2}} \\ \mathbf{L}(G_1) & \xrightarrow{\quad\mathbf{L}(\varphi)\quad} & \mathbf{L}(G_2). \end{array}$$

Proof. For $x \in \mathbf{L}(G_1)$, we consider the smooth homomorphism

$$\gamma_x \in \mathrm{Hom}(\mathbb{R}, G_1), \quad \gamma_x(t) = \exp_{G_1}(tx).$$

According to Lemma 6.2.4, we have

$$\varphi \circ \gamma_x(t) = \exp_{G_2}(ty)$$

for $y = (\varphi \circ \gamma_x)'(0) = \mathbf{L}(\varphi)x$, because $\varphi \circ \gamma_x \colon \mathbb{R} \to G_2$ is a smooth group homomorphism. For $t = 1$ we obtain in particular $\exp_{G_2}(\mathbf{L}(\varphi)x) = \varphi(\exp_{G_1}(x))$, which we had to show. $\square$

Corollary 6.2.10. *Let G_1 and G_2 be Lie groups and $\varphi\colon G_1 \to G_2$ be a group homomorphism. Then the following are equivalent:*

(a) *φ is smooth in an identity neighborhood of G_1.*

(b) *φ is smooth.*

(c) *There exists a continuous linear map $\psi\colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ satisfying*

$$\exp_{G_2} \circ \psi = \varphi \circ \exp_{G_1}. \tag{6.10}$$

Proof. (a) $\Rightarrow$ (b): Let U be an open $\mathbf{1}$-neighborhood of G_1 such that $\varphi|_U$ is smooth. Since each left translation λ_g is a diffeomorphism, $\lambda_g(U) = gU$ is an open neighborhood of g, and we have

$$\varphi(gx) = \varphi(g)\varphi(x), \quad \text{i.e.,} \quad \varphi \circ \lambda_g = \lambda_{\varphi(g)} \circ \varphi.$$

Hence the smoothness of φ on U implies the smoothness of φ on gU, and therefore that φ is smooth.

(b) $\Rightarrow$ (c): If φ is smooth, then $\psi := \mathbf{L}(\varphi)$ satisfies (6.10).

(c) $\Rightarrow$ (a): If ψ is a continuous linear map satisfying (6.10), then the fact that the exponential functions $\exp_{G_1}$ and $\exp_{G_2}$ are local diffeomorphisms, (Proposition 6.2.5) the smoothness of the linear map ψ implies (a). □

Corollary 6.2.11. *If $\varphi_1, \varphi_2 \colon G_1 \to G_2$ are morphisms of Lie groups with $\mathbf{L}(\varphi_1) = \mathbf{L}(\varphi_2)$, then φ_1 and φ_2 coincide on the identity component of G_1.*

Proof. In view of Proposition 6.2.9, we have for $x \in \mathbf{L}(G_1)$:

$$\varphi_1(\exp_{G_1}(x)) = \exp_{G_2}(\mathbf{L}(\varphi_1)x) = \exp_{G_2}(\mathbf{L}(\varphi_2)x) = \varphi_2(\exp_{G_1}(x)),$$

so that φ_1 and φ_2 coincide on the image of $\exp_{G_1}$, hence on the subgroup generated by this set. Now the assertion follows from Lemma 6.2.8. □

Proposition 6.2.12. *For a morphism $\varphi \colon G_1 \to G_2$ of Lie groups, the following assertions hold:*

(1) $\ker \mathbf{L}(\varphi) = \{x \in \mathbf{L}(G_1) \colon \exp_{G_1}(\mathbb{R}x) \subseteq \ker \varphi\}$.

(2) *φ is an open map if and only if $\mathbf{L}(\varphi)$ is surjective.*

(3) *If $\mathbf{L}(\varphi)$ and φ are bijective, then φ is an isomorphism of Lie groups.*

Proof. (1) The condition $x \in \ker \mathbf{L}(\varphi)$ is equivalent to

$$\{\mathbf{1}\} = \exp_{G_2}(\mathbb{R}\,\mathbf{L}(\varphi)x) = \varphi(\exp_{G_1}(\mathbb{R}x)).$$

(2) Suppose first that φ is an open map. Since $\exp_{G_i}$, $i = 1, 2$, are local diffeomorphisms,

$$\exp_{G_2} \circ \mathbf{L}(\varphi) = \varphi \circ \exp_{G_1} \tag{6.11}$$

implies that there exists some 0-neighborhood in $\mathbf{L}(G_1)$ on which $\mathbf{L}(\varphi)$ is an open map, hence that $\mathbf{L}(\varphi)$ is surjective.

If, conversely, $\mathbf{L}(\varphi)$ is surjective, the Open Mapping Theorem for linear operators between Banach spaces ([Ru73]) implies that $\mathbf{L}(\varphi)$ is open. Now relation (6.11) implies that there exists an open $\mathbf{1}$-neighborhood U_1 in G_1 such that $\varphi|_{U_1}$ is an open map. We claim that this implies that φ is an open map. In fact, suppose that $O \subseteq G_1$ is open and $g \in O$. Then there exists an open $\mathbf{1}$-neighborhood U_2 of G_1 with $gU_2 \subseteq O$ and $U_2 \subseteq U_1$. Then

$$\varphi(O) \supseteq \varphi(gU_2) = \varphi(g)\varphi(U_2),$$

and since $\varphi(U_2)$ is open in G_2, we see that $\varphi(O)$ is a neighborhood of $\varphi(g)$, hence that $\varphi(O)$ is open because $g \in O$ was arbitrary.

(3) From the relation $\exp_{G_2} \circ \mathbf{L}(\varphi) = \varphi \circ \exp_{G_1}$ and the bijectivity of φ we derive that the group homomorphism φ^{-1} satisfies

$$\varphi^{-1} \circ \exp_{G_2} = \exp_{G_1} \circ \mathbf{L}(\varphi)^{-1},$$

so that Corollary 6.2.10 implies that φ^{-1} is also smooth. □

The Adjoint Representation

The Lie functor associates linear automorphisms of the Lie algebra with conjugations on the Lie group. The resulting representation of the Lie group is called the adjoint representation. Its interplay with the exponential function will be important in the entire theory.

Definition 6.2.13. (a) We know that, for each Banach space V, the group $\mathrm{GL}(V)$ carries a natural Lie group structure (Example 4.1.12). For a Lie group G, a smooth homomorphism $\pi\colon G \to \mathrm{GL}(V)$ is a called a *representation of G on V* (cf. Definition 3.1.13).

Any representation defines a smooth action of G on V via

$$\sigma(g, v) := \pi(g)(v).$$

In this sense, representations are the same as *linear actions*, i.e., actions on vector spaces for which the maps σ_g are linear.

As a consequence of Proposition 6.1.4, we obtain

Proposition 6.2.14. *If $\pi\colon G \to \mathrm{GL}(V)$ is a representation of G, then $\mathbf{L}(\pi)\colon \mathbf{L}(G) \to \mathfrak{gl}(V)$ is a representation of its Lie algebra $\mathbf{L}(G)$.*

The representation $\mathbf{L}(\pi)$ obtained in Proposition 6.2.14 from the group representation π is called the *derived representation*. This is motivated by the fact that for each $x \in \mathbf{L}(G)$ we have

$$\mathbf{L}(\pi)(x) = \frac{d}{dt}\Big|_{t=0} e^{t\,\mathbf{L}(\pi)x} = \frac{d}{dt}\Big|_{t=0} \pi(\exp_G tx).$$

Let G be a Lie group and $\mathbf{L}(G)$ its Lie algebra. For $g \in G$, we recall the conjugation automorphism $c_g \in \mathrm{Aut}(G), c_g(x) = gxg^{-1}$, and define

$$\mathrm{Ad}(g) := \mathbf{L}(c_g) \in \mathrm{Aut}(\mathbf{L}(G)).$$

Then

$$\mathrm{Ad}(g_1g_2) = \mathbf{L}(c_{g_1g_2}) = \mathbf{L}(c_{g_1}) \circ \mathbf{L}(c_{g_2}) = \mathrm{Ad}(g_1)\,\mathrm{Ad}(g_2)$$

shows that $\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathbf{L}(G))$ is a group homomorphism. It is called the *adjoint representation*. To see that it is smooth, we observe that, for each $x \in \mathbf{L}(G)$, we have

$$\mathrm{Ad}(g)x = T_{\mathbf{1}}(c_g)x = T_{\mathbf{1}}(\lambda_g \circ \rho_{g^{-1}})x = T_{g^{-1}}(\lambda_g)T_{\mathbf{1}}(\rho_{g^{-1}})x = 0_g \cdot x \cdot 0_{g^{-1}}$$

in the Lie group $T(G)$ (Lemma 6.1.1). Since the multiplication in $T(G)$ is smooth, the representation Ad of G on $\mathbf{L}(G)$ is smooth (cf. Exercise 6.1.10), and

$$\mathbf{L}(\mathrm{Ad})\colon \mathbf{L}(G) \to \mathfrak{gl}(\mathbf{L}(G))$$

is a representation of $\mathbf{L}(G)$ on $\mathbf{L}(G)$. The following lemma gives a formula for this representation.

Lemma 6.2.15. $\mathbf{L}(\mathrm{Ad}) = \mathrm{ad}$, *i.e.,* $\mathbf{L}(\mathrm{Ad})(x)(y) = [x, y]$.

Proof. Let $x, y \in \mathbf{L}(G)$ and $x_l, y_l \in \mathcal{V}(G)$ be the corresponding left invariant vector fields. Corollary 6.1.6 implies for $g \in G$ the relation

$$(c_g)_* y_l = \big(\mathbf{L}(c_g)y\big)_l = \big(\mathrm{Ad}(g)y\big)_l.$$

On the other hand, the left invariance of y_l leads to

$$(c_g)_* y_l = (\rho_g^{-1} \circ \lambda_g)_* y_l = (\rho_g^{-1})_*(\lambda_g)_* y_l = (\rho_g^{-1})_* y_l.$$

Next we observe that $\Phi_t^{x_l} = \rho_{\exp_G(tx)}$ is the flow of the vector field x_l (Lemma 6.2.4), so that Theorem 5.3.12 implies that

$$[x_l, y_l] = \mathcal{L}_{x_l} y_l = \frac{d}{dt}\Big|_{t=0} (\Phi_{-t}^{x_l})_* y_l = \frac{d}{dt}\Big|_{t=0} (c_{\exp_G(tx)})_* y_l = \frac{d}{dt}\Big|_{t=0} \big(\mathrm{Ad}(\exp_G(tx))y\big)_l.$$

Evaluating in $\mathbf{1}$, we get

$$[x, y] = [x_l, y_l](\mathbf{1}) = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}(\exp_G(tx))y = \mathbf{L}(\mathrm{Ad})(x)(y). \qquad \square$$

Combining Proposition 6.2.9 with Lemma 6.2.15, we obtain the important formula

$$\mathrm{Ad} \circ \exp_G = \exp_{\mathrm{Aut}(\mathbf{L}(G))} \circ \mathrm{ad},$$

i.e.,

$$\mathrm{Ad}(\exp_G(x)) = e^{\mathrm{ad}\, x} \quad \text{for} \quad x \in \mathbf{L}(G). \tag{6.12}$$

Lemma 6.2.16. *For a Lie group G, the kernel of the adjoint representation* $\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathbf{L}(G))$*, is given by the centralizer*

$$Z_G(G_0) := \{g \in G \colon (\forall x \in G_0)\, gx = xg\},$$

of the identity component G_0. If, in addition, G is connected, then

$$\ker \mathrm{Ad} = Z(G).$$

Proof. Since G_0 is connected, the automorphism $c_g|_{G_0}$ of G_0 is trivial if and only if $\mathbf{L}(c_g) = \mathrm{Ad}(g)$ is trivial (Corollary 6.2.11). This implies the lemma. $\square$

One-parameter groups of Lie groups

With the same proof as for $\mathcal{A}^\times$, we obtain:

Lemma 6.2.17. *Let G be a Lie group, $\varepsilon > 0$ and $\gamma\colon [0, \varepsilon] \to G$ be a continuous curve with $\gamma(0) = \mathbf{1}$. If $\gamma'(0)$ exists, then*

$$\lim_{n\to\infty} \gamma\Big(\frac{1}{n}\Big)^n = \exp(\gamma'(0)).$$

Proposition 6.2.18. *If G is a Lie group and $x, y \in \mathbf{L}(G)$, then we have the Trotter Product Formula*

$$\lim_{k\to\infty}\Big(\exp_G\frac{x}{k}\exp_G\frac{y}{k}\Big)^k = \exp_G(x+y)$$

and the Commutator Formula

$$\lim_{k\to\infty}\Big(\exp_G\frac{x}{k}\exp_G\frac{y}{k}\exp_G -\frac{x}{k}\exp_G -\frac{y}{k}\Big)^{k^2} = \exp_G([x,y]).$$

Proof. To obtain the product formula, we consider the smooth curve

$$\gamma(t) := \exp_G(tx)\exp_G(ty)$$

with $\gamma(0) = \mathbf{1}$ and

$$\gamma'(0) = T_{(\mathbf{1},\mathbf{1})}(m_G)(x,y) = x + y.$$

The product formula now follows from Lemma 6.2.17.

For the commutator formula, we consider the smooth curve

$$\begin{aligned}\beta(t) &:= \exp_G(tx)\exp_G(ty)\exp_G(-tx)\exp_G(-ty) = \exp_G(t\operatorname{Ad}(\exp_G(tx)y)\exp_G(-ty)\\ &= \exp_G(te^{t\operatorname{ad}x}y)\exp_G(-ty)\end{aligned}$$

with $\beta(0) = \mathbf{1}$.

Let $U \subseteq \mathbf{L}(G)$ be an open 0-neighborhood on which $\exp_U := \exp_G|_U$ is a diffeomorphism onto an open subset of G (Proposition 6.2.5) and define

$$a * b := \exp_U^{-1}(\exp_U a \exp_U b) \quad \text{for} \quad \exp_U a \exp_U b \in \exp_G(U).$$

For sufficiently small t, we then have $\beta(t) = \exp_G(\alpha(t))$ with

$$\alpha(t) = te^{t\operatorname{ad}x}y * -ty$$

Using $T_{(\mathbf{1},\mathbf{1})}(m_G)(a,b) = a + b$ twice, we obtain with the preceding paragraph

$$\beta'(0) = (x+y) + (-x-y) = 0,$$

resp., $\alpha'(0) = 0$. We then put $\gamma(t) := \beta(\sqrt{t}) = \exp_G(\alpha(\sqrt{t}))$. In the proof of Lemma 2.2.4, we have seen that $\gamma'(0) = \frac{1}{2}\alpha''(0)$ exists. Therefore Lemma 6.2.17 implies that

$$\beta\Big(\frac{1}{k}\Big)^{k^2} = \gamma\Big(\frac{1}{k^2}\Big)^{k^2} \to \exp_G(\gamma'(0)) = \exp_G\Big(\frac{1}{2}\alpha''(0)\Big),$$

so that it remains to show that $\alpha''(0) = 2[x,y]$.

The smooth function

$$F(t,s) := se^{t\operatorname{ad}x}y * -sy$$

satisfies $F(t,0) = F(0,s) = 0$ and

$$\frac{\partial F}{\partial s}(t,0) = e^{t\operatorname{ad}x}y - y,$$

so that

$$\frac{\partial^2 F}{\partial t \partial s}(0,0) = [x,y].$$

We now derive from $\alpha(t) = F(t,t)$

$$\alpha'(t) = \frac{\partial F}{\partial t}(t,t) + \frac{\partial F}{\partial s}(t,t),$$

and thus

$$\alpha''(t) = \frac{\partial^2 F}{\partial t^2}(t,t) + 2\frac{\partial^2 F}{\partial t \partial s}(t,t) + \frac{\partial^2 F}{\partial s^2}(t,t).$$

Since $F(t,0) = F(0,s) = 0$ implies $\frac{\partial^2 F}{\partial t^2}(0,0) = \frac{\partial^2 F}{\partial s^2}(0,0) = 0$, we finally obtain

$$\alpha''(0) = 2\frac{\partial^2 F}{\partial t \partial s}(0,0) = 2[x,y].$$

□

Corollary 6.2.19. *If G is an abelian Lie group, then*

$$\exp_G(x+y) = \exp_G(x)\exp_G(y) \quad \textit{for} \quad x,y \in \mathbf{L}(G)$$

and

$$\exp_G \colon (\mathbf{L}(G), +) \to G$$

is a covering morphism of Lie groups. In particular,

$$G_0 \cong \mathbf{L}(G)/\ker \exp_G.$$

Proof. The first assertion follows from the Trotter Product Formula. It implies that $\exp_G$ is a morphism of Lie groups from the Banach space $\mathbf{L}(G)$ to G. It factors through a bijective morphism of Lie groups $\mathbf{L}(G)/\ker \exp_G \to G_0$, which is an isomorphism of Lie groups by Proposition 6.2.12(3). □

Remark 6.2.20. For finite dimensional abelian Lie groups, the preceding corollary is the key to their classification. Then $\mathbf{L}(G) \cong \mathbb{R}^n$ for some n and $\ker \exp_G$ is a discrete subgroup, hence isomorphic to $\mathbb{Z}^k$, which leads to $G_0 \cong \mathbb{T}^k \times \mathbb{R}^{n-k}$ (cf. Example 4.3.6).

Theorem 6.2.21. (One-parameter Group Theorem) *Let G be a Lie group. For each $x \in \mathbf{L}(G)$, the map $\gamma_x \colon (\mathbb{R},+) \to G, t \mapsto \exp_G(tx)$ is a smooth group homomorphism. Conversely, every continuous one-parameter group $\gamma \colon \mathbb{R} \to G$ is of this form.*

Proof. The first assertion is an immediate consequence of Lemma 6.2.4(c). It therefore remains to show that each continuous one-parameter group γ of G is a γ_x for some $x \in \mathbf{L}(G)$. Let $U = -U$ be a convex 0-neighborhood in $\mathbf{L}(G)$ for which $\exp_G|_U$ is a diffeomorphism onto an open subset of G and put $U_1 := \frac{1}{2}U$. Since γ is continuous in 0, there exists an $\varepsilon > 0$ such that $\gamma([-\varepsilon,\varepsilon]) \subseteq \exp_G(U_1)$. Then $\alpha(t) := (\exp_G|_U)^{-1}(\gamma(t))$ defines a continuous curve $\alpha \colon [-\varepsilon,\varepsilon] \to U_1$ with $\exp(\alpha(t)) = \gamma(t)$ for $|t| \leq \varepsilon$.

For any such t we then have

$$\exp_G\left(2\alpha(\tfrac{t}{2})\right) = \exp_G(\alpha(\tfrac{t}{2}))^2 = \gamma(\tfrac{t}{2})^2 = \gamma(t) = \exp_G(\alpha(t)),$$

so that the injectivity of $\exp_G$ on U yields

$$\alpha(\tfrac{t}{2}) = \tfrac{1}{2}\alpha(t) \quad \text{for} \quad |t| \leq \varepsilon.$$

Inductively we thus obtain

$$\alpha(\tfrac{t}{2^k}) = \tfrac{1}{2^k}\alpha(t) \quad \text{for} \quad |t| \leq \varepsilon, k \in \mathbb{N}. \tag{6.13}$$

In particular, we obtain

$$\alpha(t) \in \frac{1}{2^k}U_1 \quad \text{for} \quad |t| \leq \frac{\varepsilon}{2^k}.$$

For $n \in \mathbb{Z}$ with $|n| \leq 2^k$ and $|t| \leq \frac{\varepsilon}{2^k}$ we now have $|nt| \leq \varepsilon$, $n\alpha(t) \in \frac{n}{2^k}U_1 \subseteq U_1$, and

$$\exp_G(n\alpha(t)) = \gamma(t)^n = \gamma(nt) = \exp_G(\alpha(nt)).$$

Therefore the injectivity of $\exp_G$ on U_1 yields

$$\alpha(nt) = n\alpha(t) \quad \text{for} \quad n \leq 2^k, |t| \leq \frac{\varepsilon}{2^k}. \tag{6.14}$$

Combining (6.13) and (6.14), leads to

$$\alpha(\tfrac{n}{2^k}t) = \tfrac{n}{2^k}\alpha(t) \quad \text{for} \quad |t| \leq \varepsilon, k \in \mathbb{N}, |n| \leq 2^k.$$

Since the set of all numbers $\frac{nt}{2^k}$, $n \in \mathbb{Z}$, $k \in \mathbb{N}$, $|n| \leq 2^k$, is dense in the interval $[-t, t]$, the continuity of α implies that

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

In particular, α is smooth and of the form $\alpha(t) = tx$ for $x = \varepsilon^{-1}\alpha(\varepsilon)$. Hence $\gamma(t) = \exp_G(tx)$ for $|t| \leq \varepsilon$, but then $\gamma(nt) = \exp_G(ntx)$ for $n \in \mathbb{N}$ leads to $\gamma(t) = \exp_G(tx)$ for each $t \in \mathbb{R}$. □

Theorem 6.2.22. (Automatic Smoothness Theorem) *Each continuous homomorphism $\varphi\colon G \to H$ of Lie groups is smooth.*

Proof. From Theorem 6.2.21 we know that, for every Lie group G, the map

$$\mathbf{L}(G) \to \mathrm{Hom}_c(\mathbb{R}, G), \quad x \mapsto \gamma_x, \quad \gamma_x(t) := \exp_G(tx)$$

is a bijection, where $\mathrm{Hom}_c(\mathbb{R}, G)$ denotes the set of all continuous one-parameter groups of G. For $x \in \mathbf{L}(G_1)$, we consider the continuous homomorphism $\varphi \circ \gamma_x \in \mathrm{Hom}_c(\mathbb{R}, G_2)$. Since this one-parameter group is smooth (Theorem 6.2.21), it is of the form

$$\varphi \circ \gamma_x(t) = \exp_{G_2}(ty)$$

for $y = (\varphi \circ \gamma_x)'(0) \in \mathbf{L}(G_2)$. We define a map $\mathbf{L}(\varphi)\colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ by $\mathbf{L}(\varphi)x := (\varphi \circ \gamma_x)'(0)$. For $t = 1$ we then obtain

$$\exp_{G_2} \circ \mathbf{L}(\varphi) = \varphi \circ \exp_{G_1}\colon \mathbf{L}(G_1) \to G_2. \tag{6.15}$$

Next we show that $\mathbf{L}(\varphi)$ is a linear map. Our definition immediately shows that $\mathbf{L}(\varphi)\lambda x = \lambda\,\mathbf{L}(\varphi)x$ for each $x \in \mathbf{L}(G_1)$. Further, the Product Formula (Proposition 6.2.18) yields for $x, y \in \mathbf{L}(G)$:

$$\exp_{G_2}(\mathbf{L}(\varphi)(x+y)) = \varphi\big(\exp_{G_1}(x+y)\big) = \lim_{k\to\infty} \varphi\Big(\exp_{G_1}\big(\tfrac{1}{k}x\big)\exp_{G_1}\big(\tfrac{1}{k}y\big)\Big)^k$$
$$= \lim_{k\to\infty} \Big(\exp_{G_2}\big(\tfrac{1}{k}\,\mathbf{L}(\varphi)x\big)\exp_{G_2}\big(\tfrac{1}{k}\,\mathbf{L}(\varphi)y\big)\Big)^k = \exp_{G_2}\big(\mathbf{L}(\varphi)x + \mathbf{L}(\varphi)y\big).$$

This proves that $\mathbf{L}(\varphi)(x+y) = \mathbf{L}(\varphi)x + \mathbf{L}(\varphi)y$, so that $\mathbf{L}(\varphi)$ is indeed a linear map. Since $\exp_{G_2}$ is a local diffeomorphism in 0, (6.15) and the continuity of φ implies that $\mathbf{L}(\varphi)$ is continuous on some 0-neighborhood, and since it is a linear map, it is continuous (cf. Exercise 1.1.3), hence in particular smooth. Since $\exp_{G_1}$ is a local diffeomorphism in 0, (6.15) now implies that φ is smooth in an identity neighborhood of G_1, hence smooth by Corollary 6.2.10. □

Corollary 6.2.23. *A topological group G carries at most one Lie group structure.*

Proof. If G_1 and G_2 are two Lie groups which are isomorphic as topological groups, then the Automatic Smoothness Theorem applies to each topological isomorphism $\varphi\colon G_1 \to G_2$ and shows that φ is smooth. It likewise applies to φ^{-1}, so that φ is an isomorphism of Lie groups. □

The Baker–Campbell–Dynkin–Hausdorff Formula

In this subsection we show that the formula

$$\exp_G(x * y) = \exp_G x \exp_G y,$$

where $x * y$, for sufficiently small elements $x, y \in \mathfrak{g} = \mathbf{L}(G)$, is given by the Hausdorff series (cf. Proposition 2.3.6), also holds for the exponential function of a general Lie group G with Lie algebra $\mathfrak{g}$.

Definition 6.2.24. For a smooth function $f\colon M \to G$ of a smooth manifold M with values in the Lie group G, we define its *(left) logarithmic derivative* as the function

$$\delta(f)\colon TM \to \mathfrak{g}, \quad \delta(f)(v) := f(m)^{-1}.T_m(f)v \quad \text{for} \quad v \in T_m(M).$$

This map is a convenient way to describe the derivative of f in terms of a less complex structure than the tangent map $Tf\colon TM \to TG$.

Lemma 6.2.25. *For two smooth maps $f, h\colon M \to G$, the logarithmic derivative of the pointwise products fh and fh^{-1} is given by the*

(1) *Product Rule:* $\delta(fh) = \delta(h) + \mathrm{Ad}(h^{-1})\delta(f)$, *and the*

(2) *Quotient Rule:* $\delta(fh^{-1}) = \mathrm{Ad}(h)(\delta(f) - \delta(h))$.

Proof. Writing $fg = m_G \circ (f, g)$, we obtain from

$$T_{(a,b)}(m_G)(v, w) = v \cdot b + a \cdot w$$

for $a, b \in G$ and $v, w \in \mathbf{L}(G) \subseteq TG$ (Lemma 6.1.1), the relation

$$T(fh) = T(m_G) \circ (T(f), T(h)) = T(f) \cdot h + f \cdot T(h) \colon T(M) \to T(G),$$

where $f \cdot T(h)$, resp., $T(f) \cdot h$ refers to the pointwise product in the group $T(G)$, containing G as the zero section (Lemma 6.1.1). This immediately leads to the Product Rule

$$\delta(fh) = (fh)^{-1} \cdot (T(f) \cdot h + f \cdot T(h)) = h^{-1} \cdot (\delta(f) \cdot h) + \delta(h) = \mathrm{Ad}(h)^{-1}\delta(f) + \delta(h).$$

For $h = f^{-1}$, we then obtain

$$0 = \delta(ff^{-1}) = \mathrm{Ad}(f)\delta(f) + \delta(f^{-1}),$$

hence $\delta(f^{-1}) = -\mathrm{Ad}(f)\delta(f)$. This in turn leads to

$$\delta(fh^{-1}) = \mathrm{Ad}(h)\delta(f) + \delta(h^{-1}) = \mathrm{Ad}(h)\delta(f) - \mathrm{Ad}(h)\delta(h),$$

which is the Quotient Rule. □

Remark 6.2.26. For any $g \in G$ and a smooth function $f \colon M \to G$, the function $g \cdot f = \lambda_g \circ f$ has the same logarithmic derivative as f because

$$\delta(g \cdot f) = \delta(f) + \mathrm{Ad}(f)^{-1}\delta(g) = \delta(f)$$

is a consequence of the Product Rule and the fact that $\delta(g) = 0$ for the constant map with value g.

Proposition 6.2.27. *The logarithmic derivative of* $\exp_G$ *is given by*

$$\delta(\exp_G)(x) = \Phi(\mathrm{ad}\, x) \in \mathcal{L}(\mathfrak{g}), \qquad \textit{where} \qquad \Phi(z) := \frac{1 - e^{-z}}{z} = \sum_{k=1}^{\infty} \frac{(-z)^{k-1}}{k!}.$$

Proof. Fix $t, s \in \mathbb{R}$. Then the smooth functions $f, f_t, f_s \colon \mathbf{L}(G) \to G$, given by

$$f(x) := \exp_G((t + s)x), \quad f_t(x) := \exp_G(tx) \quad \text{and} \quad f_s(x) := \exp_G(sx),$$

satisfy $f = f_t f_s$ pointwise on $\mathbf{L}(G)$. The Product Rule (Lemma 6.2.25) therefore implies that

$$\delta(f) = \delta(f_s) + \mathrm{Ad}(f_s)^{-1}\delta(f_t).$$

For the smooth curve $\psi \colon \mathbb{R} \to \mathbf{L}(G), \psi(t) := \delta(\exp_G)_{tx}(ty)$, we now obtain

$$\begin{aligned} \psi(t + s) &= \delta(f)_x(y) = \delta(f_s)_x(y) + \mathrm{Ad}(f_s)^{-1}\delta(f_t)_x(y) \\ &= \psi(s) + \mathrm{Ad}(\exp_G(-sx))\psi(t). \end{aligned}$$

We have $\psi(0) = 0$ and

$$\psi'(0) = \lim_{t\to 0} \delta(\exp_G)_{tx}(y) = \delta(\exp_G)_0(y) = y,$$

so that taking derivatives with respect to t in 0, leads with (6.12) to

$$\psi'(s) = \mathrm{Ad}(\exp_G(-sx))y = e^{-\operatorname{ad}(sx)}y.$$

Now the assertion follows by integration from

$$\delta(\exp_G)_x(y) = \psi(1) = \int_0^1 \psi'(s)\,ds$$

and $\int_0^1 e^{-s\operatorname{ad} x}\,ds = \sum_{k=0}^\infty \frac{(-\operatorname{ad} x)^k}{(k+1)!} = \Phi(\operatorname{ad} x)$, which we saw already in the proof of Proposition 2.3.2. □

Let $U \subseteq \mathfrak{g}$ be a convex 0-neighborhood for which $\exp_G|_U$ is a diffeomorphism onto an open subset of G and $V \subseteq U$ a smaller convex open 0-neighborhood with $\exp_G V \exp_G V \subseteq \exp_G U$. Put $\log_U := (\exp_G|_U)^{-1}$ and define

$$x * y := \log_U(\exp_G x \exp_G y) \quad \text{for} \quad x, y \in V.$$

This defines a smooth map $V \times V \to U$. Fix $x, y \in V$. Then the smooth curve $F(t) := x * ty \in U$ satisfies $\exp_G F(t) = \exp_G(x)\exp_G(ty)$, so that the logarithmic derivative of this curve is

$$y = \delta(\exp_G)_{F(t)}F'(t) = \Phi(\operatorname{ad} F(t))F'(t).$$

We now choose U so small that the power series $\Psi(z) = \frac{z\log z}{z-1}$ from Lemma 2.3.4 satisfies

$$\Psi(e^{\operatorname{ad} z})\Phi(\operatorname{ad} z) = \mathrm{id}_{\mathfrak{g}} \quad \text{for} \quad z \in U$$

(Lemma 2.3.4). For $z = F(t)$, we then arrive with Proposition 6.2.27 at

$$F'(t) = \Psi(e^{\operatorname{ad} F(t)})y.$$

Now the same arguments as in Propositions 2.3.5 and 2.3.6 imply that

$$x * y = F(1) = x + y + \frac{1}{2}[x, y] + \cdots$$

is given by the convergent *Hausdorff series*:

Proposition 6.2.28. *If G is a Lie group, then there exists a convex 0-neighborhood $V \subseteq \mathfrak{g}$ such that for $x, y \in V$ the Hausdorff series*

$$x * y := x + \sum_{\substack{k,m\geq 0\\ p_i+q_i>0}} \frac{(-1)^k}{(k+1)(q_1+\ldots+q_k+1)} \frac{(\operatorname{ad} x)^{p_1}(\operatorname{ad} y)^{q_1}\ldots(\operatorname{ad} x)^{p_k}(\operatorname{ad} y)^{q_k}(\operatorname{ad} x)^m}{p_1!q_1!\ldots p_k!q_k!m!}y.$$

converges and satisfies

$$\exp_G(x * y) = \exp_G(x)\exp_G(y).$$

Exercises for Section 6.2

Exercise 6.2.1. Let G be a connected Lie group and $x \in \mathfrak{g} = \mathbf{L}(G)$. Show that the corresponding left invariant vector field $x_l \in \mathcal{V}(G)$ is biinvariant, i.e., also invariant under all right multiplications, if and only if $x \in \mathfrak{z}(\mathfrak{g}) := \{z \in \mathfrak{g} \colon \operatorname{ad} z = 0\}$.

Exercise 6.2.2. A vector field X on a Lie group G is called *right invariant* if for each $g \in G$, the vector field $(\rho_g)_*X = T(\rho_g) \circ X \circ \rho_g^{-1}$ coincides with X. We write $\mathcal{V}(G)^r$ for the set of right invariant vector fields on G. Show that:

(1) The evaluation map $\operatorname{ev}_{\mathbf{1}} \colon \mathcal{V}(G)^r \to T_{\mathbf{1}}(G)$ is a linear isomorphism.

(2) If X is right invariant, then there exists a unique $x \in T_{\mathbf{1}}(G)$ such that $X(g) = x_r(g) := T_{\mathbf{1}}(\rho_g)x = x \cdot 0_g$ (w.r.t. the multiplication in $T(G)$).

(3) If X is right invariant, then $\widetilde{X} := (\eta_G)_*X := T(\eta_G) \circ X \circ \eta_G^{-1}$ is left invariant and vice versa.

(4) Show that $(\eta_G)_*x_r = -x_l$ and $[x_r, y_r] = -[x, y]_r$ for $x, y \in T_{\mathbf{1}}(G)$.

(5) Show that each right invariant vector field is complete and determine its flow.

Exercise 6.2.3. No one-parameter group $\gamma : \mathbb{R} \to \mathrm{SU}_2(\mathbb{C})$ is injective, in particular, the image of $\gamma(\mathbb{R})$ is a circle group.

Exercise 6.2.4. (i) Let A be a diagonalizable endomorphism of the finite dimensional complex vector space V and let $h(z) := \sum_{n=0}^\infty a_n z^n$ be a complex power series converging on $\mathbb{C}$. We define $h(A) := \sum_{n=0}^\infty a_n A^n$ Then

$$\ker h(A) = \bigoplus_{z \in h^{-1}(0) \cap \operatorname{Spec}(A)} \ker(A - z\mathbf{1}).$$

(ii) Let A be an endomorphism of the real vector space V for which the complex linear extension $A_{\mathbb{C}} \colon V_{\mathbb{C}} \to V_{\mathbb{C}}$ is diagonalizable. Then $\operatorname{Spec}(A) := \operatorname{Spec}(A_{\mathbb{C}})$ decomposes into the subsets

$$S_{\mathrm{re}} := \operatorname{Spec}(A) \cap \mathbb{R} \quad \text{and} \quad S_{\mathrm{im}} := \operatorname{Spec}(A) \setminus S_{\mathrm{re}}.$$

Let h be as above and assume, in addition, that $h(\overline{z}) = \overline{h(z)}$. Then $h(A)V \subseteq V$ and

$$\ker h(A) = \bigoplus_{z \in h^{-1}(0) \cap S_{\mathrm{re}}} \ker(A - z\mathbf{1}) \oplus \bigoplus_{x+iy \in h^{-1}(0) \cap S_{\mathrm{im}}, y>0} \ker\big(A^2 - 2xA + (x^2 + y^2)\big).$$

Exercise 6.2.5. Let $A \in \operatorname{End}(V)$, where V is a finite dimensional real vector space and

$$\Phi(z) := \frac{1 - e^{-z}}{z} = \sum_{k=0}^\infty \frac{(-1)^k z^k}{(k+1)!}.$$

Show that $\Phi(A)$ is invertible if and only if

$$\operatorname{Spec}(A) \cap 2\pi i\mathbb{Z} \subseteq \{0\}.$$

Exercise 6.2.6. (Divisible groups) An abelian group D is called *divisible* if for each $d \in D$ and $n \in \mathbb{N}$ there exists an $a \in D$ with $a^n = d$. Show that:

(1)* If G is an abelian group, H a subgroup and $f : H \to D$ a homomorphism into an abelian divisible group D, then there exists an extension of f to a homomorphism $\widetilde{f} : G \to D$.

(2) If G is an abelian group and D a divisible subgroup, then $G \cong D \times H$ for some subgroup H of G.

Exercise 6.2.7. (Non-connected abelian Lie groups) Let A be an abelian Lie group. Show that:

(1) If $\dim A < \infty$, then the identity component of A_0 is isomorphic to $\mathbb{R}^k \times \mathbb{T}^m$ for some $k, m \in \mathbb{N}_0$.

(2) A_0 is divisible (cf. Corollary 6.2.19).

(3) $A \cong A_0 \times \pi_0(A)$, where $\pi_0(A) := A/A_0$ (Exercise 6.2.6).

(4) There exists a discrete abelian group D with $A \cong A_0 \times D$.

Chapter 7

Subgroups of Lie Groups

We have seen in Corollary 6.2.23 that a topological group carries at most one Lie group structure. Therefore we call a subgroup H of a Lie group G a Lie subgroup if it carries a Lie group structure compatible with the subspace topology. In this chapter we take a closer look at this concept. In particular, we show that Lie subgroups are always closed and that, for finite dimensional Lie groups, the converse is also true. This makes it easy to find Lie group structures on all closed matrix groups.

7.1 Lie Subgroups

Definition 7.1.1. Let H be a subgroup of a Lie group G. It is a topological group with respect to the subspace topology. We call H a *Lie subgroup* if it carries a Lie group structure compatible with this topology. According to Corollary 6.2.23, this Lie group structure is unique if it exists.

Remark 7.1.2. If $H \subseteq G$ is a Lie subgroup and $i_H \colon H \to G$ the inclusion map, then i_H is a continuous homomorphism of Lie groups, hence smooth by the Automatic Smoothness Theorem and therefore a morphism of Lie groups.

Definition 7.1.3. Let G be a Lie group and $H \leq G$ be a ubgroup. We define the set

$$\mathbf{L}^e(H) := \{x \in \mathbf{L}(G) \colon \exp_G(\mathbb{R}x) \subseteq H\}$$

and observe that $\mathbb{R}\,\mathbf{L}^e(H) \subseteq \mathbf{L}^e(H)$ follows immediately from the definition.

Proposition 7.1.4. *If $H \leq G$ is a closed subgroup of the Lie group G, then $\mathbf{L}^e(H)$ is a closed real subalgebra of $\mathbf{L}(G)$.*

Proof. With the Product and Commutator Formula for general Lie groups (Proposition 6.2.18), the arguments are the same as for linear groups (Proposition 3.1.4). □

Lemma 7.1.5. *Every Lie subgroup H of a Lie group G is closed.*

Proof. Let $i_H: H \to G$ be the embedding map, which is a morphism of Lie groups. In view of

$$i_H \circ \exp_H = \exp_G \circ \mathbf{L}(i_H) \tag{7.1}$$

and the fact that $\exp_H$ and $\exp_G$ are local homeomorphisms, the bijective continuous linear map $\mathbf{L}(i_H): \mathbf{L}(H) \to \mathbf{L}(i_H)\mathbf{L}(H)$ also is a local homeomorphism, hence a homeomorphism by (cf. Exercise 4.1.2). In particular, $\mathrm{im}(\mathbf{L}(i_H))$ is a complete subspace of $\mathbf{L}(G)$, hence closed. That it coincides with $\mathbf{L}^e(H)$ follows immediately from the correspondence of one-parameter subgroups of H with elements of $\mathbf{L}^e(H)$, resp., elements of $\mathbf{L}(H)$ (Lemma 6.2.4). We may therefore identify the Lie algebra $\mathbf{L}(H)$ of the Lie group H with the subset $\mathbf{L}^e(H)$ of $\mathbf{L}(G)$. In this sense we then have

$$\exp_H = \exp_G |_{\mathbf{L}(H)}.$$

Let $V_G \subseteq \mathbf{L}(G)$ be a 0-neighborhood for which $\exp_G |_{V_G}$ is a diffeomorphism onto an open subset of G and $V_H \subseteq \mathbf{L}(H) \cap V_G$ be a 0-neighborhood for which $\exp_H |_{V_H}$ is a diffeomorphism onto an open subset of H. Since $\exp_G |_{V_G}$ is a homeomorphism, there exists an open 0-neighborhood $W \subseteq V_G \subseteq \mathbf{L}(G)$ with

$$\exp_G(W) \cap H = \exp_H(V_H) = \exp_G(V_H).$$

Now $\exp_G(W \cap \mathbf{L}(H)) \subseteq H$ and $V_H \subseteq W \cap \mathbf{L}(H)$ lead to $W \cap \mathbf{L}(H) = V_H$. In particular V_H is closed in W because $\mathbf{L}(H)$ is closed in $\mathbf{L}(G)$, and therefore $\exp_H(V_H) = H \cap \exp_G(W)$ is closed in $\exp_G(W)$. We conclude that H is locally closed, and therefore closed (Lemma 1.2.2(ii)). □

Theorem 7.1.6. (Integral Subgroup Theorem; general version) *Let G be a Lie group and $\mathfrak{h} \subseteq \mathbf{L}(G)$ be a closed Lie subalgebra. Then the subgroup $H := \langle \exp \mathfrak{h} \rangle$ of G generated by* $\exp(\mathfrak{h})$ *carries a Lie group structure for which there exists an open 0-neighborhood $V \subseteq \mathfrak{h}$ on which the Dynkin series converges,*

$$\exp: \mathfrak{h} \to H, \quad x \mapsto \exp x$$

is smooth and maps V diffeomorphically onto an open subset of H and satisfies

$$\exp(x * y) = \exp(x)\exp(y) \quad \textit{for} \quad x, y \in V.$$

Proof. Let $U = -U \subseteq \mathbf{L}(G)$ be an open convex 0-neighborhood such that the Dynkin series $x * y$ converges for $x, y \in U$, defines a smooth map $U \times U \to \mathbf{L}(G)$, and satisfies

$$\exp_G(x * y) = \exp_G(x)\exp_G(y).$$

(Proposition 6.2.28). We may further assume that $\exp_G |_U$ is a diffeomorphism onto an open subset of G. Now the remaining arguments are the same as in the linear case (Theorem 4.2.4). □

The following proposition is a generalization of the Linear Lie Group Theorem to general Lie groups.

Proposition 7.1.7. *For a closed subgroup H of a Lie group G, the following are equivalent:*

(i) *H is a Lie subgroup.*

(ii) *There exists an open 0-neighborhood $W \subseteq \mathbf{L}^e(H)$ for which $\exp_G|_W$ is a homeomorphism onto an open $\mathbf{1}$-neighborhood in H.*

(iii) *There exists an open 0-neighborhood $V \subseteq \mathbf{L}(G)$ with $\exp_G^{-1}(H) \cap V \subseteq \mathbf{L}^e(H)$.*

Proof. (i) $\Rightarrow$ (ii): Take $U := V_H$ in the proof of Lemma 7.1.5.

(ii) $\Rightarrow$ (iii): Let $V_0 \subseteq \mathbf{L}(G)$ be an open 0-neighborhood for which $\exp_G|_{V_0}$ is a diffeomorphism onto an open subset of G. Then (ii) still holds for the smaller 0-neighborhood $W_0 := W \cap V_0$ in $\mathbf{L}^e(H)$. Since $\exp_G(W_0)$ is open in H, resp., $H \cap \exp_G(V_0)$, and $\exp_G|_{V_0}$ is a homeomorphism, there exists an open 0-neighborhood $V_1 \subseteq V_0$ with $\exp_G(V_1) \cap H = \exp_G(W_0)$. Then the injectivity of $\exp_G|_{V_1}$ implies

$$\exp_G^{-1}(H) \cap V_1 = (\exp_G|_{V_1})^{-1}(H) = W_0 \subseteq \mathbf{L}^e(H).$$

(iii) $\Rightarrow$ (i): Shrinking V, we may w.l.o.g. assume that $\exp_G|_V$ is a diffeomorphism onto an open subset of G. For $W := V \cap \mathbf{L}^e(H)$ we then obtain $\exp_G(V) \cap H = \exp_G(W)$, so that $\varphi := \exp_G|_W$ is a homeomorphism onto an open $\mathbf{1}$-neighborhood in H. Then $\exp(\mathbf{L}^e(H)) \subseteq H$ is a connected $\mathbf{1}$-neighborhood, so that $\langle \exp \mathbf{L}^e(H) \rangle$ is an open connected subgroup of H, hence coincides with the identity component H_0 of H (cf. Lemmas 1.2.2(iv) and 1.2.10). We endow H_0 with the Lie group structure from Theorem 7.1.6, and the remaining arguments are the same as in the proof of the Linear Lie Group Theorem 4.2.6. ☐

Proposition 7.1.8. *If $\varphi \colon G_1 \to G_2$ is a morphism of Lie groups and $H_2 \subseteq G_2$ is a Lie subgroup, then $H_1 := \varphi^{-1}(H_2)$ is a Lie subgroup with Lie algebra*

$$\mathbf{L}^e(H_1) = \mathbf{L}(\varphi)^{-1}\big(\mathbf{L}^e(H_2)\big).$$

In particular, $\ker \varphi$ *is a Lie subgroup of G_1 with Lie algebra* $\ker \mathbf{L}(\varphi)$.

Proof. Let $V_2 \subseteq \mathbf{L}(G_2)$ be an open 0-neighborhood with $\exp_{G_2}^{-1}(H_2) \cap V_2 \subseteq \mathbf{L}^e(H_2)$ (Proposition 7.1.7) and note that $V_1 := \mathbf{L}(\varphi)^{-1}(V_2)$ is an open 0-neighborhood in $\mathbf{L}(G_1)$.

For $x \in V_1$ with $\exp_{G_1} x \in H_1$, we then have $\exp_{G_2}(\mathbf{L}(\varphi)x) = \varphi(\exp_{G_1} x) \in \varphi(H_1) \subseteq H_2$, so that $\mathbf{L}(\varphi)x \in V_2$ implies $\mathbf{L}(\varphi)x \in \mathbf{L}^e(H_2)$, hence $x \in \mathbf{L}(\varphi)^{-1}(\mathbf{L}^e(H_2)) = \mathbf{L}^e(H_1)$ (Exercise!). We conclude that

$$\exp_{G_1}^{-1}(H_1) \cap V_1 \subseteq \mathbf{L}^e(H_1),$$

so that Proposition 7.1.7 implies that H_1 is a Lie subgroup of G_1.

Since $\{\mathbf{1}\}$ is a Lie subgroup of G_2, we see in particular that $\ker \varphi$ is a Lie subgroup of G_1. ☐

Applying the preceding proposition to the adjoint representation $\mathrm{Ad} \colon G \to \mathrm{Aut}(\mathfrak{g})$ of a connected Lie group G, we obtain in particular for $Z(G) = \ker \mathrm{Ad}$:

Corollary 7.1.9. *For every connected Lie group G, the center $Z(G)$ is a Lie subgroup.*

7.2 The Closed Subgroup Theorem

We now address more detailed information on closed subgroups of finite dimensional Lie groups. We start with three key lemmas providing the main information for the proof of the Closed Subgroup Theorem.

Lemma 7.2.1. *Let $W \subseteq \mathbf{L}(G)$ be an open 0-neighborhood for which $\exp_G|_W$ is a diffeomorphism. Further, let $H \subseteq G$ be a closed subgroup and $(y_k)_{k\in\mathbb{N}}$ be a sequence in W such that $y_k \to 0$ and $g_k := \exp_G y_k \in H \setminus \{\mathbf{1}\}$ for all $k \in \mathbb{N}$. Fix a norm $\|\cdot\|$ on $\mathbf{L}(G)$. Then every cluster point of the sequence $\left\{\frac{y_k}{\|y_k\|} : k \in \mathbb{N}\right\}$ is contained in $\mathbf{L}^e(H)$.*

Proof. Let x be such a cluster point. Replacing the original sequence by a subsequence, we may assume that we have in $\mathbf{L}(G)$:

$$x_k := \frac{y_k}{\|y_k\|} \to x.$$

Note that this implies $\|x\| = 1$. Let $t \in \mathbb{R}$ and put $p_k := \frac{t}{\|y_k\|}$. Then $tx_k = p_k y_k$, so that $y_k \to 0$ leads to

$$\exp_G(tx) = \lim_{k\to\infty} \exp_G(tx_k) = \lim_{k\to\infty} \exp_G(p_k y_k)$$

and

$$\exp_G(p_k y_k) = \exp_G(y_k)^{[p_k]} \exp_G\big((p_k - [p_k])y_k\big),$$

where $[p_k] = \max\{\ell \in \mathbb{Z} \colon \ell \le p_k\}$ is the *Gauß function.* We then have

$$\|(p_k - [p_k])y_k\| \le \|y_k\| \to 0$$

and

$$\exp_G(tx) = \lim_{k\to\infty} (\exp_G y_k)^{[p_k]} = \lim_{k\to\infty} g_k^{[p_k]} \in H,$$

because H is closed. This implies $x \in \mathbf{L}^e(H)$. ☐

Lemma 7.2.2. *Let $H \subseteq G$ be a closed subgroup and $E \subseteq \mathbf{L}(G)$ be a finite dimensional vector subspace complementing $\mathbf{L}(H)$. Then there exists a 0-neighborhood $U_E \subseteq E$ with*

$$H \cap \exp_G(U_E) = \{\mathbf{1}\}.$$

Proof. We argue by contradiction. If a neighborhood U_E with the required properties does not exist, then for each compact convex 0-neighborhood $V_E \subseteq E$, we have for each $k \in \mathbb{N}$:

$$(\exp_G \tfrac{1}{k} V_E) \cap H \neq \{\mathbf{1}\}.$$

For each $k \in \mathbb{N}$ we therefore find $y_k \in V_E$ with $\mathbf{1} \neq g_k := \exp_G(\frac{y_k}{k}) \in H$. Now the compactness of V_E implies that the sequence $(y_k)_{k\in\mathbb{N}}$ is bounded, so that $\frac{y_k}{k} \to 0$, which implies $g_k \to \mathbf{1}$. Now let $x \in E$ be a cluster point of the sequence $\frac{y_k}{\|y_k\|}$ which lies in the compact subset $S_E := \{z \in E \colon \|z\| = 1\}$ of the finite dimensional normed space E. According to Lemma 7.2.1, we have $x \in \mathbf{L}^e(H) \cap E = \{0\}$ because $g_k \in H \cap W$ for k sufficiently large. We arrive at a contradiction to $\|x\| = 1$. This proves the lemma. ☐

Lemma 7.2.3. *Suppose that* $\dim G < \infty$ *and that* $E, F \subseteq \mathbf{L}(G)$ *are subspaces with* $E \oplus F = \mathbf{L}(G)$. *Then the map*

$$\Phi\colon E \times F \to G, \quad (x,y) \mapsto \exp_G(x)\exp_G(y),$$

restricts to a diffeomorphism of a neighborhood of $(0,0)$ *to an open* $\mathbf{1}$*-neighborhood in* G.

Proof. The Chain Rule implies that

$$\begin{aligned} T_{(0,0)}(\Phi)(x,y) &= T_{(\mathbf{1},\mathbf{1})}(m_G) \circ (T_0(\exp_G)x, T_0(\exp_G)y) \\ &= T_{(\mathbf{1},\mathbf{1})}(m_G)(x,y) = x + y, \end{aligned}$$

Since the addition map $E \times F \to \mathbf{L}(G) \cong T_{\mathbf{1}}(G)$ is bijective, the Inverse Function Theorem implies that Φ restricts to a diffeomorphism of an open neighborhood of $(0,0)$ in $E \times F$ onto an open neighborhood of $\mathbf{1}$ in G. □

Theorem 7.2.4. (Closed Subgroup Theorem) *Every closed subgroup of a finite dimensional Lie group* G *is a Lie subgroup.*

Proof. Let $H \subseteq G$ be a closed subgroup and $E \subseteq \mathbf{L}(G)$ be a vector space complement of the subspace $\mathbf{L}^e(H)$ of $\mathbf{L}(G)$. We define

$$\Phi\colon E \times \mathbf{L}^e(H) \to G, \quad (x,y) \mapsto \exp_G x \exp_G y.$$

According to Lemma 7.2.3, there exist open 0-neighborhoods $U_E \subseteq E$ and $U_H \subseteq \mathbf{L}^e(H)$ such that

$$\Phi_1 := \Phi|_{U_E \times U_H} : U_E \times U_H \to \exp_G(U_E)\exp_G(U_H)$$

is a diffeomorphism onto an open $\mathbf{1}$-neighborhood in G. In view of Lemma 7.2.2, we may even choose U_E so small that $\exp_G(U_E) \cap H = \{\mathbf{1}\}$.

Since $\exp_G(U_H) \subseteq H$, the condition

$$g = \exp_G x \exp_G y \in H \cap (\exp_G(U_E)\exp_G(U_H))$$

implies $\exp_G x = g(\exp_G y)^{-1} \in H \cap \exp_G U_E = \{\mathbf{1}\}$. Therefore

$$H \supseteq \exp_G(U_H) = H \cap (\exp_G(U_E)\exp_G(U_H))$$

is an open $\mathbf{1}$-neighborhood in H. In view of Proposition 7.1.7, this completes the proof. □

Example 7.2.5. We take a closer look at closed subgroups of the Lie group $(V,+)$, where V is a finite-dimensional vector space. From Example 6.2.3 we know that $\exp_V = \mathrm{id}_V$. Let $H \subseteq V$ be a closed subgroup. Then

$$\mathbf{L}^e(H) = \{x \in V \colon \mathbb{R}x \subseteq H\} \subseteq H$$

is the largest vector subspace contained in H. Let $E \subseteq V$ be a vector space complement for $\mathbf{L}^e(H)$. Then $V \cong \mathbf{L}^e(H) \times E$, and we derive from $\mathbf{L}^e(H) \subseteq H$ that

$$H \cong \mathbf{L}^e(H) \times (E \cap H).$$

Lemma 7.2.2 implies the existence of some 0-neighborhood $U_E \subseteq E$ with $U_E \cap H = \{0\}$, hence that $H \cap E$ is discrete because 0 is an isolated point of $H \cap E$. Now Exercise 1.2.7 implies the existence of linearly independent elements $f_1, \ldots, f_k \in E$ with

$$E \cap H = \mathbb{Z}f_1 + \ldots + \mathbb{Z}f_k.$$

We conclude that

$$H \cong \mathbf{L}^e(H) \times \mathbb{Z}^k \cong \mathbb{R}^d \times \mathbb{Z}^k \quad \text{for} \quad d = \dim \mathbf{L}^e(H).$$

Note that $\mathbf{L}^e(H)$ coincides with the identity component H_0 of H.

Example 7.2.6. How bad closed subgroups can be is illustrated by the following example due to K. H. Hofmann: We consider the real Hilbert space $G := L^2([0,1],\mathbb{R})$ as a Banach–Lie group. Then the subgroup $H := L^2([0,1],\mathbb{Z})$ of all those functions which almost everywhere take values in $\mathbb{Z}$ is a closed subgroup. Since the one-parameter subgroups of G are of the form $\mathbb{R}f$, $f \in G$, we have $\mathbf{L}^e(H) = \{0\}$. On the other hand, the group H is arcwise connected and even contractible because the map $F\colon [0,1] \times H \to H$ given by

$$F(t,f)(x) := \begin{cases} f(x) & \text{for } 0 \leq x \leq t \\ 0 & \text{for } t < x \leq 1 \end{cases}$$

is continuous with $F(1,f) = f$ and $F(0,f) = 0$. We conclude that the closed subgroup H of G is NOT a Lie subgroup.

This pathology can be avoided by the assumption that the subgroup is connected by C^1 arcs.

Example 7.2.7. (Closed Subgroups of $\mathbb{T}$) Let $H \subseteq \mathbb{T} \subseteq (\mathbb{C}^\times, \cdot)$ be a closed proper (=different from $\mathbb{T}$) subgroup. Since $\dim \mathbb{T} = 1$, it follows that $\mathbf{L}^e(H) = \{0\}$, so that the Identity Neighborhood Theorem implies that H is discrete, hence finite because $\mathbb{T}$ is compact.

If $q\colon \mathbb{R} \to \mathbb{T}$ is the covering projection, $q^{-1}(H)$ is a closed proper subgroup of $\mathbb{R}$, hence cyclic, which implies that $H = q(q^{-1}(H))$ is also cyclic. Therefore H is one of the groups

$$C_n = \{z \in \mathbb{T}\colon z^n = 1\}$$

of n-th roots of unity.

Example 7.2.8. (Subgroups of $\mathbb{T}^2$) (a) Let $H \subseteq \mathbb{T}^2$ be a closed proper subgroup. Then $\mathbf{L}^e(H) \neq \mathbf{L}(\mathbb{T}^2)$ implies $\dim H < \dim \mathbb{T}^2 = 2$. Further, H is compact, so that the group $\pi_0(H)$ of connected components of H is finite.

If $\dim H = 0$, then H is finite, and for $n := |H|$ it is contained in a subgroup of the form $C_n \times C_n$, where $C_n \subseteq \mathbb{T}$ is the subgroup of n-th roots of unity (cf. Example 7.2.7).

If $\dim H = 1$, then H_0 is a compact connected 1-dimensional Lie group, hence isomorphic to $\mathbb{T}$ (cf. Examples 4.3.6). Therefore $H_0 = \exp_{\mathbb{T}^2}(\mathbb{R}x)$ for some $x \in \mathbf{L}^e(H)$ with $\exp_{\mathbb{T}^2}(x) = (e^{2\pi i x_1}, e^{2\pi i x_2}) = (1,1)$, which is equivalent to $x \in \mathbb{Z}^2$. We conclude that the Lie algebras of the closed subgroups are of the form $\mathbf{L}^e(H) = \mathbb{R}x$ for some $x \in \mathbb{Z}^2$.

(b) For each $\theta \in \mathbb{R} \setminus \mathbb{Q}$ the image of the 1-parameter group

$$\gamma \colon \mathbb{R} \to \mathbb{T}^2, \quad t \mapsto (e^{i\theta t}, e^{it})$$

is not closed because γ is injective. Hence the closure of $\gamma(\mathbb{R})$ is a closed subgroup of dimension at least 2, which shows that $\gamma(\mathbb{R})$ is dense in $\mathbb{T}^2$. The subgroup $\gamma(\mathbb{R})$ is called a dense wind.

Exercises for Section 7.2

Exercise 7.2.1. Show that a linear subspace $E \subseteq \mathbb{R}^d \cong \mathbf{L}(\mathbb{T}^d)$ is the Lie algebra $\mathbf{L}^e(H)$ of a closed subgroup $H \subseteq \mathbb{T}^d$ if and only if it is *rational*, i.e., spanned by $E \cap \mathbb{Q}^d$, resp., $E \cap \mathbb{Z}^d$.

Exercise 7.2.2. (Torus complements) Show that, for every subgroup $H \cong \mathbb{T}^k$ of $G := \mathbb{T}^d$ there exists another subgroup $C \cong \mathbb{T}^{d-k}$ with $G \cong H \times C$. Hint: Put $\Gamma := \ker \exp_G \cong \mathbb{Z}^d$. Verify that $\Gamma_H := \Gamma \cap \mathbf{L}^e(H) \cong \mathbb{Z}^k$ and argue that there is a subgroup $\mathbb{Z}^{d-k} \cong \Gamma_C \subseteq \Gamma$ with $\Gamma \cong \Gamma_H \oplus \Gamma_C$. Then consider $C := \exp_G(\mathrm{span}\Gamma_C)$.

Exercise 7.2.3. (Closed subgroups of $\mathbb{T}^d$) Show that every closed subgroup $H \subseteq \mathbb{T}^d$ is isomorphic to a product $\mathbb{T}^k \times F$ for a finite subgroup F.

Exercise 7.2.4. Let $x \in \mathbb{R}^n = \mathbf{L}(\mathbb{T}^n)$. Show that the one-parameter group $\exp(\mathbb{R}x)$ is dense in $\mathbb{T}^n$ if and only if $x_1 \neq 0$ and the real numbers $(1, x_2/x_1, \ldots, x_n/x_1)$ are linearly independent over $\mathbb{Q}$. Hint: Use Exercise 7.2.1.

Chapter 8

Integration of Lie Algebra Homomorphisms

To round off the picture of Lie groups and their Lie algebras presented in this lecture, we still have to provide the link between Lie algebras and covering groups. The main point is that, in general, one cannot integrate morphisms of Lie algebras $\mathbf{L}(G) \to \mathbf{L}(H)$ to morphisms of connected Lie groups $G \to H$ if G is not simply connected.

8.1 The Monodromy Principle and its Applications

Proposition 8.1.1. (Monodromy Principle) *Let G be a simply connected Lie group and H a group. Let V be an open symmetric connected identity neighborhood in G and $f\colon V \to H$ a function with*

$$f(xy) = f(x)f(y) \quad \text{for} \quad x, y, xy \in V.$$

Then there exists a unique group homomorphism extending f. If, in addition, H is a Lie group and f is smooth, then its extension is also smooth.

Proof. We consider the group $G \times H$ and the subgroup $S \subseteq G \times H$ generated by the subset $U := \{(x, f(x))\colon x \in V\}$. We endow U with the topology for which $x \mapsto (x, f(x)), V \to U$ is a homeomorphism. Then U is connected because V is connected. Note that $f(\mathbf{1})^2 = f(\mathbf{1}^2) = f(\mathbf{1})$ implies $f(\mathbf{1}) = \mathbf{1}$, which further leads to $\mathbf{1} = f(xx^{-1}) = f(x)f(x^{-1})$, so that $f(x^{-1}) = f(x)^{-1}$. Hence $U = U^{-1}$.

To obtain a group topology on S, we now apply Lemma 1.4.5, and observe that S is generated by U, and that (T1/2) directly follow from the corresponding properties of V and $(x, f(x))(y, f(y)) = (xy, f(xy))$ for $x, y, xy \in V$. This leads to a group topology on S, for which S is a connected topological group. Indeed, its connectedness follows from $S = \bigcup_{n\in\mathbb{N}} U^n$ and the connectedness of all sets U^n (Exercise 1.2.1). The projection $p_G\colon G \times H \to G$ induces a covering homomorphism $q\colon S \to G$ because its restriction to the open $\mathbf{1}$-neighborhood U is a homeomorphism (Exercise C.2.2(c)), and the connectedness of S and the simple connectedness of G imply that q is a homeomorphism (Corollary C.2.8). Now $F := p_H \circ q^{-1} : G \to H$ provides the required extension of f. In fact, for $x \in U$ we have $q^{-1}(x) = (x, f(x))$, and therefore $F(x) = f(x)$.

If, in addition, H is Lie and f is smooth, then the smoothness of the extension follows directly from Corollary 6.2.10. □

Theorem 8.1.2. (Integrability Theorem for Lie Algebra Homomorphisms) *Let G be a connected simply connected Lie group, H a Lie group and $\psi\colon \mathbf{L}(G) \to \mathbf{L}(H)$ a continuous Lie algebra morphism. Then there exists a unique morphism $\varphi\colon G \to H$ with $\mathbf{L}(\varphi) = \psi$.*

Proof. Let $U \subseteq \mathbf{L}(G)$ be an open connected symmetric 0-neighborhood on which the BCDH-product is defined and satisfies $\exp_G(x * y) = \exp_G(x)\exp_G(y)$ and

$$\exp_H(\psi(x) * \psi(y)) = \exp_H(\psi(x))\exp_H(\psi(y)) \quad \text{for} \quad x, y \in U$$

(Proposition 6.2.28). Assume further that V is an open 0-neighborhood with $U*U \subseteq V$ for which $\exp_G|_V$ is a homeomorphism onto an open subset of G (cf. Proposition 6.2.5) (this can always achieved by shrinking U if necessary).

The continuity of ψ and the fact that ψ is a Lie algebra homomorphism imply that for $x, y \in U$ the element $\psi(x * y)$ coincides with the convergent Hausdorff series $\psi(x) * \psi(y)$ (cf. Proposition 6.2.28). We define

$$f\colon \exp_G(U) \to H, \quad f(\exp_G(x)) := \exp_H(\psi(x)).$$

For $x, y \in U$ with $\exp_G x \exp_G y \in \exp_G(U)$ we then find some $z \in U$ with $\exp_G z = \exp_G x \exp_G y = \exp_G(x * y)$, so that the injectivity of $\exp_G$ on $V \supseteq U * U$ leads to $x * y = z \in U$. We now obtain

$$\begin{aligned} & f(\exp_G(x)\exp_G(y)) = f(\exp_G(x * y)) = \exp_H(\psi(x * y)) \\ = \; & \exp_H(\psi(x) * \psi(y)) = \exp_H(\psi(x))\exp_H(\psi(y)) = f(\exp_G(x))f(\exp_G(y)). \end{aligned}$$

Then $f\colon \exp(U) \to H$ satisfies the assumptions of Proposition 8.1.1, and we see that f extends uniquely to a group homomorphism $\varphi\colon G \to H$. Since $\exp_G$ is a local diffeomorphism, f is smooth in a **1**-neighborhood, and therefore φ is smooth. We finally observe that φ is uniquely determined by $\mathbf{L}(\varphi) = \psi$ because G is connected (Corollary 6.2.11). □

The following corollary can be viewed as an integrability condition for ψ.

Corollary 8.1.3. *If G is a connected Lie group and H is a Lie group, then, for a Lie algebra morphism $\psi\colon \mathbf{L}(G) \to \mathbf{L}(H)$, there exists a morphism $\varphi\colon G \to H$ with $\mathbf{L}(\varphi) = \psi$ if and only if $\pi_1(G) \cong \ker q_G \subseteq \ker \widetilde{\varphi}$, where $q_G\colon \widetilde{G} \to G$ is the universal covering map and $\widetilde{\varphi}\colon \widetilde{G} \to H$ is the unique morphism with $\mathbf{L}(\widetilde{\varphi}) = \psi \circ \mathbf{L}(q_G)$.*

Proof. If φ exists, then

$$(\varphi \circ q_G) \circ \exp_{\widetilde{G}} = \varphi \circ \exp_G \circ \mathbf{L}(q_G) = \exp_H \circ \psi \circ \mathbf{L}(q_G)$$

and the uniqueness of $\widetilde{\varphi}$ imply that $\widetilde{\varphi} = \varphi \circ q_G$ and hence that $\ker q_G \subseteq \ker \widetilde{\varphi}$.

If, conversely, $\ker q_G \subseteq \ker \widetilde{\varphi}$, then $\varphi(q_G(g)) := \widetilde{\varphi}(g)$ defines a continuous morphism $G \cong \widetilde{G}/\ker q_G \to H$ with $\varphi \circ q_G = \widetilde{\varphi}$ (Exercise 1.1.9) and

$$\varphi \circ \exp_G \circ \mathbf{L}(q_G) = \varphi \circ q_G \circ \exp_{\widetilde{G}} = \widetilde{\varphi} \circ \exp_{\widetilde{G}} = \exp_H \circ \psi \circ \mathbf{L}(q_G). \qquad \square$$

The following corollary is a partial converse to Corollary 6.1.6, which asserts that isomorphic Lie groups have isomorphic Lie algebras.

Corollary 8.1.4. *If G and H are simply connected Lie groups with isomorphic Lie algebras, then G and H are isomorphic.*

Proof. Let $\alpha\colon \mathbf{L}(G) \to \mathbf{L}(H)$ be an isomorphism and use Theorem 8.1.2 to find a morphism $\varphi\colon G \to H$ of Lie groups with $\mathbf{L}(\varphi) = \alpha$. We likewise find a morphism of Lie groups $\psi\colon H \to G$ with $\mathbf{L}(\psi) = \alpha^{-1}$, and then the relations

$$\mathbf{L}(\varphi \circ \psi) = \mathrm{id}_{\mathbf{L}(H)} \quad \text{and} \quad \mathbf{L}(\psi \circ \varphi) = \mathrm{id}_{\mathbf{L}(G)}$$

imply that $\varphi \circ \psi = \mathrm{id}_H$ and $\psi \circ \varphi = \mathrm{id}_G$ (Corollary 6.2.11). Therefore φ is an isomorphism of Lie groups. $\square$

Corollary 8.1.5. *If G is a simply connected Lie group with Lie algebra $\mathfrak{g}$, then the map*

$$\mathbf{L}\colon \mathrm{Aut}(G) \to \mathrm{Aut}(\mathfrak{g})$$

is an isomorphism of groups.

Proof. First, we recall from Corollary 6.1.6 that for each automorphism $\varphi \in \mathrm{Aut}(G)$ the endomorphism $\mathbf{L}(\varphi)$ of $\mathfrak{g}$ also is an automorphism. That $\mathbf{L}$ is injective follows from the connectedness of G (Corollary 6.2.11) and that $\mathbf{L}$ is surjective from the Integrability Theorem 8.1.2. $\square$

8.2 Classification of Lie Groups with given Lie Algebra

Let G and H be linear Lie groups. If $\varphi\colon G \to H$ is an isomorphism, then the functoriality of $\mathbf{L}$ directly implies that $\mathbf{L}(\varphi)\colon \mathbf{L}(G) \to \mathbf{L}(H)$ is an isomorphism (Corollary 6.1.6). In this subsection we ask to which extent a Lie group G is determined by its Lie algebra $\mathbf{L}(G)$.

Proposition 8.2.1. *A surjective morphism $\varphi\colon G \to H$ of Lie groups is a covering if and only if $\mathbf{L}(\varphi)\colon \mathbf{L}(G) \to \mathbf{L}(H)$ is a linear isomorphism.*

Proof. In view of Exercise 4.3.10, φ is a covering if and only if φ is open with discrete kernel.

In Proposition 6.2.12 we have seen that φ is open if and only if $\mathbf{L}(\varphi)$ is surjective. Since $\ker \varphi$ is a Lie subgroup by Proposition 7.1.8, it is discrete if and only if $\mathbf{L}(\ker \varphi) = \ker \mathbf{L}(\varphi) = \{0\}$, which means that $\mathbf{L}(\varphi)$ is injective. Combining these observations, we see that $\mathbf{L}(\varphi)$ is bijective if and only if φ is open with discrete kernel, i.e., a covering. $\square$

Proposition 8.2.2. *For a covering* $q\colon G_1 \to G_2$ *of connected Lie groups, the following equalities hold*

$$q\big(Z(G_1)\big) = Z(G_2) \quad \text{and} \quad Z(G_1) = q^{-1}(Z(G_2)).$$

Proof. Since q is a covering, $\mathbf{L}(q)\colon \mathbf{L}(G_1) \to \mathbf{L}(G_2)$ is an isomorphism of Lie algebras, and the adjoint representations satisfy

$$\mathrm{Ad}_{G_2}(q(g_1)) \circ \mathbf{L}(q) = \mathbf{L}(q) \circ \mathrm{Ad}_{G_1}(g_1).$$

Hence

$$Z(G_1) = \ker \mathrm{Ad}_{G_1} = q^{-1} \ker \mathrm{Ad}_{G_2} = q^{-1}(Z(G_2)).$$

Now the claim follows from the surjectivity of q. □

Theorem 8.2.3. *Two connected Lie groups G and H have isomorphic Lie algebras if and only if their universal covering groups $\widetilde{G}$ and $\widetilde{H}$ are isomorphic.*

Proof. If $\widetilde{G}$ and $\widetilde{H}$ are isomorphic, then we clearly have

$$\mathbf{L}(G) \cong \mathbf{L}(\widetilde{G}) \cong \mathbf{L}(\widetilde{H}) \cong \mathbf{L}(H)$$

(Proposition 8.2.1 and Corollary 6.1.6).

Conversely, Corollary 8.1.4 shows that any isomorphism $\psi\colon \mathbf{L}(G) \to \mathbf{L}(H)$ leads to an isomorphism $\widetilde{G} \to \widetilde{H}$. □

Combining the preceding theorem with Theorem 4.3.5, we obtain:

Corollary 8.2.4. *Let G be a connected Lie group and $q_G\colon \widetilde{G} \to G$ be the universal covering morphism of connected Lie groups. Then, for each discrete central subgroup $\Gamma \subseteq \widetilde{G}$, the group $\widetilde{G}/\Gamma$ is a connected Lie group with $\mathbf{L}(\widetilde{G}/\Gamma) \cong \mathbf{L}(G)$ and, conversely, each Lie group with the same Lie algebra as G is isomorphic to some quotient $\widetilde{G}/\Gamma$.*

Example 8.2.5. We now describe a pair of nonisomorphic Lie groups with isomorphism Lie algebras and isomorphic fundamental groups.

Let

$$\widetilde{G} := \mathrm{SU}_2(\mathbb{C}) \times \mathrm{SU}_2(\mathbb{C})$$

whose center is $C_2 \times C_2$,

$$G := \widetilde{G}/(C_2 \times \{\mathbf{1}\}) \cong \mathrm{SO}_3(\mathbb{R}) \times \mathrm{SU}_2(\mathbb{C})$$

and

$$H := \widetilde{G}/\{(\mathbf{1}, \mathbf{1}), (-\mathbf{1}, -\mathbf{1})\} \cong \mathrm{SO}_4(\mathbb{R}),$$

where the latter isomorphy can be obtained by considered $\mathrm{SO}_4(\mathbb{R})$ as a group acting on the skew field $\mathbb{H}$ of quaternions. Then $\pi_1(G) \cong \pi_1(H) \cong C_2$ (Theorem 4.3.5), but there is no automorphism of $\widetilde{G}$ mapping $\pi_1(G)$ to $\pi_1(H)$.

Indeed, one can show that the two direct factors are the only nontrivial connected normal subgroups of $\widetilde{G}$, so that each automorphism of $\widetilde{G}$ either preserves both or exchanges them. Since $\pi_1(H)$ is not contained in any of them, it cannot be mapped to $\pi_1(G)$ by an automorphism of $\widetilde{G}$.

Examples 8.2.6. Here are some examples of pairs of linear Lie groups with isomorphic Lie algebras:

(1) $G = \mathrm{SO}_3(\mathbb{R})$ and $\widetilde{G} \cong \mathrm{SU}_2(\mathbb{C})$ (Example 4.3.8).

(2) $G = \mathrm{SO}_{2,1}(\mathbb{R})_0$ and $H = \mathrm{SL}_2(\mathbb{R})$: In this case we actually have a covering morphism $\varphi\colon H \to G$ coming from the adjoint representation

$$\mathrm{Ad}\colon \mathrm{SL}_2(\mathbb{R}) \to \mathrm{GL}(\mathbf{L}(H)) \cong \mathrm{GL}_3(\mathbb{R}).$$

On $\mathbf{L}(H) = \mathfrak{sl}_2(\mathbb{R})$ we consider the symmetric bilinear form given by $\beta(x,y) := \frac{1}{2}\operatorname{tr}(xy)$ and the basis

$$e_1 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e_2 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad e_3 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Then the matrix B of β with respect to this basis is

$$B := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

One easily verifies that

$$\operatorname{Im}\mathrm{Ad} \subseteq \mathrm{O}(\mathbf{L}(H), \beta) \cong \mathrm{O}_{2,1}(\mathbb{R}),$$

and since $\mathrm{ad}\colon \mathbf{L}(H) \to \mathfrak{o}_{2,1}(\mathbb{R})$ is injective between spaces of the same dimension 3 (Exercise), it is bijective. Therefore $\operatorname{im}\mathrm{Ad} = \langle \exp \mathfrak{o}_{2,1}(\mathbb{R})\rangle = \mathrm{SO}_{2,1}(\mathbb{R})_0$ and Proposition 8.2.1 imply that

$$\mathrm{Ad}\colon \mathrm{SL}_2(\mathbb{R}) \to \mathrm{SO}_{2,1}(\mathbb{R})_0$$

is a covering morphism. Its kernel is given by $Z(\mathrm{SL}_2(\mathbb{R})) = \{\pm\mathbf{1}\}$ (Lemma 6.2.16).

One can show that both groups are homeomorphic to $\mathbb{T} \times \mathbb{R}^2$, and topologically the map Ad is like $(z,x,y) \mapsto (z^2,x,y)$, a two-fold covering.

Example 8.2.7. Let $G = \mathrm{SL}_2(\mathbb{R})$ and $H = \mathrm{SO}_{2,1}(\mathbb{R})_0$ and recall that $\widetilde{G} \cong \widetilde{H}$ follows from $\mathfrak{sl}_2(\mathbb{R}) \cong \mathfrak{so}_{2,1}(\mathbb{R})$ (cf. Example 8.2.6).

We further have $q_G(Z(\widetilde{G})) \subseteq Z(G) = \{\pm\mathbf{1}\}$ and $\pi_1(G) = \ker q_G \subseteq Z(\widetilde{G})$ (cf. Proposition 8.2.2). Likewise $q_H(Z(\widetilde{G})) \subseteq Z(H) = \{\mathbf{1}\}$ implies

$$Z(\widetilde{G}) \cong \pi_1(H) \cong \pi_1(\mathrm{O}_2(\mathbb{R}) \times \mathrm{O}_1(\mathbb{R})) \cong \mathbb{Z},$$

where the latter is a consequence of the polar decomposition. This implies that $Z(\widetilde{G}) \cong \mathbb{Z}$, where

$$\pi_1(G) \cong 2\mathbb{Z} \quad \text{and} \quad \pi_1(H) \cong \mathbb{Z} = Z(\widetilde{G}).$$

Therefore G and H are not isomorphic, but they have isomorphic Lie algebras and isomorphic fundamental groups.

Exercises for Section 8.2

Exercise 8.2.1. Let G be a connected linear Lie group. Show that

$$\mathbf{L}(Z(G)) = \mathfrak{z}(\mathbf{L}(G)) := \{x \in \mathbf{L}(G)\colon (\forall y \in \mathbf{L}(G))\ [x,y] = 0\}.$$

Appendix A

Basic Topological Concepts

In this appendix we collect some basic notions concerning topological spaces.

A.1 Topological Spaces

Definition A.1.1. Let X be a set. We write $\mathbb{P}(X)$ for the *power set of* X, i.e., the set of all subsets of X. A subset $\tau \subseteq \mathbb{P}(X)$, whose elements are called *open sets*, is called a *topology on* X if the following axioms are satisfied:

(T1) $\emptyset, X$ are open sets.

(T2) Finite intersections of open sets are open.

(T3) Arbitrary unions of open sets are open.

If τ is a topology on X, then the pair (X, τ) is called a topological space.[1] To simplify notation we often write X instead of (X, τ) for a topological space whose underlying set is X.

Example A.1.2. (a) If (X, d) is a metric space, then we call a subset $O \subseteq X$ *open* if for each $x \in O$ there exists an $\varepsilon > 0$ with

$$B_\varepsilon(x) := \{y \in X : d(x, y) < \varepsilon\} \subseteq O.$$

Then the system τ_d of open subsets of X is a topology and the triangle inequality immediately implies that the balls $B_\varepsilon(x)$ are open. We call it the *topology defined (or induced) by the metric d on X.*

(b) $\tau = \{X, \emptyset\}$ is a topology on X, called the *chaotic topology.*

(c) $\tau = \mathbb{P}(X)$ is a topology on X, called the *discrete topology.* In this case (X, τ) is called a *discrete space.*

---

[1] Metric spaces were first studied by Maurice Fréchet in 1906 and topological spaces were introduced later in 1914 by Felix Hausdorff (1868–1942). It is interesting to observe that the more abstract notion of a topological spaces was conceived later.

Definition A.1.3. Let (X, τ) be a topological space.

(a) A subset $C \subseteq X$ is called *closed* if its complement $X \setminus C$ is open.

(b) For $x \in X$ we call a subset $U \subseteq X$ a *neighborhood of* x if there exists an open subset $O \subseteq X$ with $x \in O \subseteq U$. We write $\mathfrak{U}(x)$, or $\mathfrak{U}_X(x)$, for the set of all neighborhoods of x.

Lemma A.1.4. *If (X, τ) is a topological space, then the set of all closed subsets of X has the following properties:*

(C1) $\emptyset, X$ *are closed.*

(C2) *Finite unions of closed sets are closed.*

(C3) *Arbitrary intersections of closed sets are closed.*

Proof. This follows immediately from (O1)-(O3) by taking complements and using de Morgan's Rules: $(\bigcup_{i \in I} O_i)^c = \bigcap_{i \in I} O_i^c$ and $(\bigcap_{i \in I} O_i)^c = \bigcup_{i \in I} O_i^c$. □

Definition A.1.5. Let (X, τ) be a topological space and $E \subseteq X$ a subset.

(a) $\overline{E} := \bigcap \{F \subseteq X : E \subseteq F, F \text{ closed}\}$ is called the *closure of* E. This is the smallest closed subset of X containing E.

(b) $E^0 := \bigcup \{U \subseteq X : U \subseteq E, U \text{ open}\}$ is called the *interior of* E. This is the largest open subset contained in E.

(c) $\partial E := \overline{E} \setminus E^0$ is called the *boundary of* E.

Lemma A.1.6. *Let (X, τ) be a topological space, $E \subseteq X$ and $x \in X$. Then the following assertions hold:*

(1) $x \in E^0 \Leftrightarrow (\exists U \in \mathfrak{U}(x))\ U \subseteq E \Leftrightarrow E \in \mathfrak{U}(x)$.

(2) $x \in \overline{E} \Leftrightarrow (\forall U \in \mathfrak{U}(x))\ U \cap E \neq \emptyset$.

(3) $x \in \partial E \Leftrightarrow (\forall U \in \mathfrak{U}(x))\ U \cap E \neq \emptyset$ *and* $U \not\subseteq E$.

Definition A.1.7. A topological space (X, τ) is called a *Hausdorff space* or *separated* if for each pair (x, y) of different points in X there exist disjoint neighborhoods of x and y.

Remark A.1.8. (a) Metric spaces are Hausdorff spaces because for $x \neq y$ and $2r < d(x, y)$ the open balls $B_r(x)$ and $B_r(y)$ are disjoint.

(b) Let X be a Hausdorff space and $x \in X$. Then

$$\{x\} = \bigcap \mathfrak{U}_X(x).$$

Moreover, the one point set $\{x\}$ is closed because its complement $\{x\}^c = \bigcup_{y \neq x} B_{d(x,y)}(y)$ is open.

A.2 Continuous maps

After introducing the concept of a topological space as a pair (X, τ) of a set X with a distinguished collection of subsets called open, we now explain what the corresponding structure preserving maps are. They are called continuous maps, resp., functions.

Definition A.2.1. Let (X, τ_X) and (Y, τ_Y) be topological spaces.

(a) A map $f\colon X \to Y$ is called *continuous* if for each open subset $O \subseteq Y$ the inverse image $f^{-1}(O)$ is an open subset of X. Then f is also called a *morphism of topological space.*

We write $C(X, Y)$ for the set of continuous maps $f\colon X \to Y$.

(b) A continuous map $f\colon X \to Y$ is called a *homeomorphism* or *topological isomorphism* if there exists a continuous map $g\colon Y \to X$ with

$$f \circ g = \mathrm{id}_Y \quad \text{and} \quad g \circ f = \mathrm{id}_X .$$

(c) A map $f\colon X \to Y$ is said to be *open* if for each open subset $O \subseteq X$, the image $f(O)$ is an open subset of Y. We similarly define *closed* maps $f\colon X \to Y$ as those mapping closed subsets of X to closed subsets of Y.

Proposition A.2.2. *If $f : X \to Y$ and $g : Y \to Z$ are continuous maps, then their composition $g \circ f : X \to Z$ is continuous.*

Proof. For any open subset $O \subseteq Z$, the set $(g \circ f)^{-1}(O) = f^{-1}(g^{-1}(O))$ is open in X because $g^{-1}(O)$ is open in Y. □

Lemma A.2.3. *(a) If $f\colon X \to Z$ is a continuous map and $Y \subseteq X$ a subset, then $f|_Y\colon Y \to Z$ is continuous with respect to the subspace topology on Y.*

(b) If $f\colon X \to Z$ is a map and $Y \subseteq Z$ is a subset containing $f(X)$, then f is continuous if and only if the corestriction $f^{|Y}\colon X \to Y$ is continuous with respect to the subspace topology on Y.

Proof. (a) If $O \subseteq Z$ is open, then $(f|_Y)^{-1}(O) = f^{-1}(O) \cap Y$ is open in the subspace topology. Therefore $f|_Y$ is continuous.

(b) For a subset $O \subseteq Z$, we have

$$f^{-1}(O) = f^{-1}(O \cap Y) = (f^{|Y})^{-1}(O \cap Y).$$

This implies that f is continuous if and only if the corestriction $f^{|Y}$ is continuous. □

Presently, we only have a global concept of continuity. To define also what it means that a function is continuous in a point, we use the concept of a neighborhood.

Definition A.2.4. Let X and Y be topological spaces and $x \in X$. A function $f\colon X \to Y$ is said to be *continuous in* x if for each neighborhood V of $f(x)$ there exists a neighborhood U of x with $f(U) \subseteq V$. Note that this condition is equivalent to $f^{-1}(V)$ being a neighborhood of x.

Remark A.2.5. If (X, d_X) and (Y, d_Y) are metric spaces, then a map $f\colon X \to Y$ is continuous in $x \in X$ if and only if

$$(\forall \varepsilon > 0)(\exists \delta > 0) \quad f\big(B_\delta(x)\big) \subseteq B_\varepsilon\big(f(x)\big).$$

This follows easily from the observation that $V \subseteq Y$ is a neighborhood of $f(x)$ if and only if it contains some ball $B_\varepsilon(f(x))$ and $U \subseteq X$ is a neighborhood of x if and only if it contains some ball $B_\delta(x)$.

Lemma A.2.6. *Let $f\colon X \to Y$ and $g\colon Y \to Z$ be maps between topological spaces. If f is continuous in x and g is continuous in $f(x)$, then the composition $g \circ f$ is continuous in x.*

Proof. Let V be a neighborhood of $g(f(x))$ in Z. Then the continuity of g in $f(x)$ implies the existence of a neighborhood V' of $f(x)$ with $g(V') \subseteq V$. Further, the continuity of f in x implies the existence of a neighborhood U of x in X with $f(U) \subseteq V'$, and then $(g \circ f)(U) \subseteq g(V') \subseteq V$. Therefore $g \circ f$ is continuous in x. □

Proposition A.2.7. *For a map $f\colon X \to Y$ between topological spaces, the following are equivalent:*

(1) *f is continuous.*

(2) *f is continuous in each $x \in X$.*

(3) *Inverse images of closed subsets of Y under f are closed.*

(4) *For each subset $M \subseteq X$, we have $f(\overline{M}) \subseteq \overline{f(M)}$.*

Proof. (1) ⇒ (2): Let $V \subseteq Y$ be a neighborhood of $f(x)$. Then the continuity of f implies that $U := f^{-1}(V^0)$ is an open subset of X containing x, hence a neighborhood of x with $f(U) \subseteq V$.

(2) ⇒ (1): Let $O \subseteq Y$ be open and $x \in f^{-1}(O)$. Since f is continuous in x, $f^{-1}(O)$ is a neighborhood of x, and since x is arbitrary, the set $f^{-1}(O)$ is open.

(1) ⇔ (3): If $A \subseteq Y$ is closed, then $f^{-1}(A) = f^{-1}(A^c)^c$ implies that all these subsets of X are closed if and only if all sets $f^{-1}(A^c)$ are open, which is equivalent to the continuity of f.

(3) ⇒ (4): The inverse image $f^{-1}(\,\overline{f(M)}\,)$ is a closed subset of X containing M, hence also $\overline{M}$.

(4) ⇒ (3): If $A \subseteq Y$ is closed and $M := f^{-1}(A)$, then $f(\overline{M}) \subseteq \overline{f(M)} \subseteq A$ implies that $\overline{M} \subseteq M$, i.e., M is closed. □

Proposition A.2.8. *For a continuous map $f\colon X \to Y$, the following are equivalent:*

(1) *f is a homeomorphism.*

(2) *f is bijective and $f^{-1}\colon Y \to X$ is continuous.*

(3) *f is bijective and open.*

(4) *f is bijective and closed.*

Proof. (1) $\Leftrightarrow$ (2): Let $g\colon Y \to X$ be continuous with $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$. Then f is bijective, and $f^{-1} = g$ is continuous.

If, conversely, f is bijective and f^{-1} is continuous, then we see with $g := f^{-1}$ that f is a homeomorphism.

(2) $\Leftrightarrow$ (3): For $O \subseteq X$ we have $f(O) = (f^{-1})^{-1}(O)$. That this set if open for each open subset $O \subseteq X$ is equivalent to f being open and to f^{-1} being continuous.

(2) $\Leftrightarrow$ (4): For $A \subseteq X$ we have $f(A) = (f^{-1})^{-1}(A)$. That this set if closed for each closed subset $A \subseteq X$ is equivalent to f being closed and to f^{-1} being continuous (Proposition A.2.7). ☐

A.3 Creating new topologies

Definition A.3.1. For a subset $\mathcal{A} \subseteq \mathbb{P}(X)$ the set

$$\tau := \langle \mathcal{A} \rangle_{top} := \bigcap \{\sigma \colon \mathcal{A} \subseteq \sigma, \sigma \text{ topology }\}$$

is a topology on X; the coarsest topology on X for which all sets in $\mathcal{A}$ are open. Therefore τ is called the *topology generated by* $\mathcal{A}$ and $\mathcal{A}$ is called a *subbasis of the topology* τ. The set $\mathcal{A}$ is called as *basis* of the topology τ if every set in τ is a union of sets in $\mathcal{A}$.

Definition A.3.2. (a) Let $\sim$ be an equivalence relation on the topological space (X, τ), $Y := X/\sim = \{[x] \colon x \in X\}$ the set of equivalence classes, and $q\colon X \to Y, x \mapsto [x]$ the quotient map. Then

$$\sigma := \{U \subseteq Y \colon q^{-1}(U) \in \tau\}$$

is a topology on Y called the *quotient topology*.

The quotient topology has the property that a map $f\colon Y \to Z$ into a topological space Z is continuous if and only if the map $f \circ q\colon X \to Z$ is continuous (Exercise). In particular, $q\colon X \to Y$ is continuous.

(b) Let $(X_1, \tau_1), \ldots, (X_n, \tau_n)$ be topological spaces and $X := X_1 \times \ldots \times X_n$ the Cartesian product set. Then the system of all subsets $U \subseteq X$ of the form

$$U = U_1 \times \ldots \times U_n, \quad U_j \in \tau_j,$$

is a basis for a topology on X called the *product topology*.

The product topology has the property that a map

$$f\colon Z \to X, \quad z \mapsto f(z) = (f_1(z), \ldots, f_n(z))$$

is continuous if and only if all maps $f_j\colon Z \to X_j$ are continuous. In particular, the projections $p_j\colon X \to X_j$ are continuous maps.

(c) If (X, τ) is a topological space and $Y \subseteq X$ a subset, then

$$\tau_Y := \{U \cap Y \colon U \in \tau\}$$

is a topology on Y called the *subspace topology*.

The subspace topology has the property that a map $f\colon Z \to Y$ is continuous if and only if the corresponding map $f_X\colon Z \to X, z \mapsto f(z)$ is continuous. In particular, the inclusion map $\iota_Y\colon Y \to X$ is continuous.

A.4 Compactness

Definition A.4.1. A topological space (X, τ) is said to be *compact* if it is separated and each open cover, i.e., each collection $(U_j)_{j\in J}$ of open subsets of X with $\bigcup_{j\in J} U_j = X$, contains a finite subcover. This means that there exist $j_1, \ldots, j_n \in J$ with

$$X = U_{j_1} \cup \ldots \cup U_{j_n}.$$

Lemma A.4.2. (Compactness of subspaces) (a) *If (X, τ) is a separated space and $Y \subseteq X$, then Y is compact with respect to the subspace topology if and only if each open cover of Y by open subsets of X contains a finite subcover.*

(b) *If X is separated and $Y \subseteq X$ is compact, then Y is closed in X. If, conversely, X is compact and $Y \subseteq X$ closed, then Y is compact.*

Proof. (b) Suppose first that Y is a compact subspace of the separated space X. Let $(U_i)_{i\in I}$ be an open covering of Y and pick open subsets $O_i \subseteq X$ with $O_i \cap Y = U_i$. Then the open subset Y^c, together with the O_i, $i \in I$, form an open covering in X. Hence there exists a finite subcovering, and this implies the existence of a finite subset $F \subseteq I$ with $Y \subseteq \bigcup_{i\in F} U_i$.

Now suppose that Y is a closed subspace of the compact space X. Let $x \in Y^c$. For each $y \in Y$ we then have $y \neq x$, and since X is separated, there exists an open subset U_y of X and an open subset V_c of X with $y \in U_y$, $x \in V_y$ and $U_y \cap V_y = \emptyset$. Then we obtain an open covering $(U_y \cap Y)_{y\in Y}$ of Y. Let $U_{y_1} \cap Y, \ldots, U_{y_n} \cap Y$ be a finite subcovering and $V := \bigcap_{i=1}^n V_{y_i}$. Then V intersects $\bigcup_{i=1}^n U_{y_i} \supseteq Y$ trivially, and therefore $x \notin \overline{Y}$. This proves that Y is closed. □

Proposition A.4.3. (a) *If X is compact, Y separated and $f\colon X \to Y$ continuous, then $f(X)$ is a compact subset of Y.*

(b) *If $f\colon X \to \mathbb{R}$ is continuous and X compact, then f is bounded and takes a minimal and a maximal value.*

Proof. (a) Let $(U_i)_{i\in I}$ be a covering of $f(X)$ by open subsets of Y. Then $(f^{-1}(U_i))_{i\in I}$ is an open covering of X, so that there exists a finite subset $F \subseteq I$ with $X \subseteq \bigcup_{i\in F} f^{-1}(U_i)$. Then

$$f(X) \subseteq \bigcup_{i\in F} f(f^{-1}(U_i)) \subseteq \bigcup_{i\in F} U_i.$$

Since $f(X)$ is separated, it follows that $f(X)$ is compact.

(b) follows from (a). □

Proposition A.4.4. *If $f\colon X \to Y$ is bijective, Y separated and X compact, then f is a homeomorphism.*

Proof. Let $A \subseteq X$ be a closed subset. Then A is compact by Lemma A.4.2. Therefore $f(A) \subseteq Y$ is compact, hence closed by Lemma A.4.2. Since f is continuous, $A \subseteq X$ is closed if and only if $f(A) \subseteq Y$ is closed, so that f is a homeomorphism. □

Corollary A.4.5. *Let $f\colon X \to Y$ be a surjective continuous map, X compact and Y separated. We define an equivalence relation $\sim$ on X by $x \sim y$ if $f(x) = f(y)$ and endow the space $X/\sim$ with the quotient topology. Then the map*

$$\overline{f}\colon X/\sim \to Y, \quad [x] \mapsto f(x)$$

is a homeomorphism.

Proof. Let $q\colon X \to X/\sim$ denote the quotient map. For $[x_1] \neq [x_2]$ we have $f(x_1) \neq f(x_2)$. Let $U_j \in \mathfrak{U}(f(x_j))$ be disjoint open neighborhoods. Then $f^{-1}(U_j)$, $j = 1, 2$, are disjoint open subsets of X. Moreover, the set $V_j := q(f^{-1}(U_j)) \subseteq X/\sim$ satisfy $q^{-1}(V_j) = f^{-1}(U_j)$. Therefore the sets V_j are disjoint open neighborhoods of the $[x_j]$. Hence $X/\sim$ is separated. Now Proposition A.4.3(a) implies that $X/\sim$ is compact, and Proposition A.4.4 applies to $\overline{f}$. □

Lemma A.4.6. *Let X and Y be topological spaces, $K_X \subseteq X$ and $K_Y \subseteq Y$ compact, and $O_j \subseteq X \times Y$, $j \in J$, open sets with $K_X \times K_Y \subseteq \bigcup_{j\in J} O_j$. Then there exist open subsets $U_X \subseteq X$, $U_Y \subseteq Y$ and $j_1, \ldots, j_n \in J$ with*

$$K_X \times K_Y \subseteq U_X \times U_Y \subseteq O_{j_1} \cup \ldots \cup O_{j_n}.$$

Proof. For each pair $(x, y) \in K_X \times K_Y$ there exists a $j(x, y) \in J$ with $(x, y) \in O_{j(x,y)}$, so that the definition of the product topology implies the existence of open neighborhoods $U_{x,y}$ of x and $V_{x,y}$ of y with $U_{x,y} \times V_{x,y} \subseteq O_{j(x,y)}$. Fix $x \in K_X$. Then the sets $(V_{x,y})_{y\in K_Y}$, form an open cover of K_Y, hence have a finite subcover $V_{x,y_1}, \ldots, V_{x,y_n}$. Let

$$U_x := \bigcap_{j=1}^{n} U_{x,y_j} \quad \text{and} \quad V_x := \bigcup_{j=1}^{n} V_{x,y_j}.$$

Then U_x and V_x are open with

$$U_x \times V_x \subseteq U \quad \text{and} \quad K_Y \subseteq V_x.$$

Now the sets $(U_x)_{x\in K_X}$ form an open cover of K_X, and we find $x_1, \ldots, x_m \in K_X$ with

$$K_X \subseteq U_{x_1} \cup \ldots \cup U_{x_m}.$$

Now we set

$$U := U_{x_1} \cup \ldots \cup U_{x_m} \supseteq K_X \quad \text{and} \quad V := V_{x_1} \cap \ldots \cap V_{x_m} \supseteq K_Y$$

and obtain $K_X \times K_Y \subseteq U \times V$, where the set $U \times V$ is contained in finitely many sets of the form $U_x \times V_x$, which in turn is contained in the union of the sets $U_{x,y_i} \times V_{x,y_i} \subseteq \mathcal{O}_{j(x,y_i)}$. We conclude that $U \times V$ is contained in a union of finitely many of the sets $\mathcal{O}_j$. □

Corollary A.4.7. *Let X and Y be topological spaces, $K_X \subseteq X$ and $K_Y \subseteq Y$ compact, and $O \subseteq X \times Y$ open with $K_X \times K_Y \subseteq O$. Then there exist open subsets $U_X \subseteq X$ and $U_Y \subseteq Y$ with*

$$K_X \times K_Y \subseteq U_X \times U_Y \subseteq O.$$

Proposition A.4.8. *If $X_1, \ldots, X_n$ are compact spaces, then their product*

$$X_1 \times \ldots \times X_n$$

is compact.

Proof. It suffices to prove the assertion for $n = 2$ and then apply induction. We apply Lemma A.4.6 with $X = K_X = X_1$, $Y = K_Y = Y_1$ and an open covering $\mathcal{O}_j$, $j \in J$, of $X \times Y$. Then Lemma A.4.6 implies the existence of a finite subcovering. □

A.5 Connectedness and arc connectedness

Definition A.5.1. Let X be a topological space. An *arc* in X is a continuous map $\gamma\colon [a, b] \to X$, where $[a, b] \subseteq \mathbb{R}$ is a compact interval. We also say that γ is an arc from $\gamma(a)$ to $\gamma(b)$.

For $p, q \in X$ we define $x \sim_a y$ if there exists an arc from p to q. We claim that $\sim_a$ defines an equivalence relation. The constant arc connects x to x, so that $x \sim_a x$ for each $x \in X$. If $x \sim_a y$ and $\gamma\colon [a, b] \to X$ connects x to y, then

$$\widetilde{\gamma}\colon [0, 1] \to X, \quad t \mapsto \gamma(b + t(a - b))$$

connects y to x, and we get $y \sim_a x$. For the transitivity, assume that $\gamma\colon [a, b] \to X$ connects x to y and that $\eta\colon [c, d] \to X$ connects y to z. Then

$$\xi\colon [0, 2] \to X, \quad \xi(t) := \begin{cases} \gamma(a + t(b - a)) & \text{for } t \in [0, 1] \\ \eta(c + (t - 1)(d - c)) & \text{for} t \in [1, 2] \end{cases}$$

is an arc connecting x to z. Therefore $x \sim_a z$.

The equivalence classes of $\sim_a$ are denoted $C_a(x)$ and called *arc components* of the topological space X. The space X is said to be *arcwise connected* if any two points of X can be connected by an arc.

Definition A.5.2. Let X be a topological space. We say that X is *connected* if $\emptyset$ and X are the only subsets of X which are closed and open. This is equivalent to saying that if $A, B \subseteq X$ are two disjoint open subsets with $X = A \cup B$, then one of them is empty.

Lemma A.5.3. (a) *Each interval in $\mathbb{R}$ is connected.*

(b) *If X is arcwise connected, then X is connected.*

(c) *If $f\colon X \to Y$ is continuous and X is connected, resp., arcwise connected, then the same holds for $f(X)$.*

Proof. (a) Let $I \subseteq \mathbb{R}$ be an interval and suppose that it is not connected. Then there exist two disjoint proper open subsets $A, B \subseteq I$ with $I = A \cup B$. Let $a \in A$ and $b \in B$. Then the compact interval $C := [a, b]$ is contained in I and $C \cap A$ is an open subset of C not containing b. Therefore $s := \sup(A \cap C) < b$. Since $A \cap C = C \setminus B$ is closed in C, we have $s \in A$. On the other hand $A \cap C$ is open in C, so that it contains a neighborhood of s, contradicting the definition of s.

(b) Let $A, B \subseteq X$ be two disjoint open subsets with $X = A \cup B$. If both are non-empty, we pick $a \in A$ and $b \in B$. Let $\gamma\colon [0,1] \to X$ be an arc from a to b. Then $[0,1] = \gamma^{-1}(A)\dot\cup\gamma^{-1}(B)$ is a decomposition into two proper open disjoint subsets. This contradicts the connectedness of $[0,1]$.

(c) Suppose first that X is connected. If $f(X)$ is not connected, then there exist two open subsets $A, B \subseteq Y$ such that $f(X) = (f(X) \cap A)\dot\cup(f(X) \cap B)$, where both are proper subsets. Then $X = f^{-1}(A) \cup f^{-1}(B)$ is a decomposition into two disjoint proper open subsets, contradicting the connectedness of X.

If X is arcwise connected and $\gamma\colon [0,1] \to X$ is an arc from x to y, then

$$f \circ \gamma\colon [0,1] \to Y$$

is an arc from $f(x) \to f(y)$. Therefore $f(X)$ is arcwise connected. □

Lemma A.5.4. *Let X be a topological space.*

(i) *If $(A_j)_{j\in J}$ are connected subspaces of X with $\bigcap_{j\in J} A_j \neq \emptyset$, then the subset $A := \bigcup_{j\in J} A_j$ of X is connected.*

(ii) *For each connected subspace $A \subseteq X$ its closure $\overline{A}$ is also connected.*

Proof. (i) Since the subspace topologies on A_j inherited from X and A are the same, we may w.l.o.g. assume that $A = X$. Suppose that $X = U_1\dot\cup U_2$, where U_1 and U_2 are open subsets. Pick $a \in \bigcap_{j\in J} A_j$ and let $k \in \{1,2\}$ with $a \in U_k$. Then for each $j \in J$ we have a disjoint decomposition

$$A_j = (A_j \cap U_1)\dot\cup(A_j \cap U_2)$$

into two open subsets of A_j. Since A_j is connected and $a \in A_j \cap U_k$, it follows that $A_j = A_j \cap U_k$, and therefore $A_j \subseteq U_k$. Thus $X = \bigcup_{j\in J} A_j = U_k$.

(ii) As above, we may w.l.o.g. assume that $X = \overline{A}$. Suppose that U_1 and U_2 are open subsets of X with $X = U_1\dot\cup U_2$. Then we obtain the disjoint decomposition

$$A = (A \cap U_1)\dot\cup(A \cap U_2),$$

and since A is connected, there exists a $k \in \{1,2\}$ with $A \subseteq U_k$. Since the complement of U_k is open, the set U_k is closed, so that $X = \overline{A} \subseteq U_k$. □

Definition A.5.5. Let X be a topological space. Then each one element subset $\{x\} \subseteq X$ is connected. Therefore

$$C(x) := \bigcup\{A \subseteq X : x \in A, A \text{ connected}\}$$

is a connected subset by Lemma A.5.4(i), and by Lemma A.5.4(ii) it is closed. It is called the *connected component of* x. It is the largest connected subset of X containing x. It follows directly from Lemma A.5.4 that the connected components are pairwise disjoint closed subsets of X.

The space X is connected if and only if $C(x) = X$ for each $x \in X$. It is called *totally disconnected* if $C(x) = \{x\}$ for each $x \in X$.

We say that X is *locally connected* if each $x \in X$ has a connected neighborhood. This implies in particular that, for each $x \in X$, the connected component $C(x)$ is a neighborhood of x. Hence the connected components of X are open subsets.

Examples A.5.6. (a) The set $[0,1] \cup [2,3]$ is locally connected but not connected.

(b) For two elements x, y in the euclidean plane $\mathbb{R}^2$ we write $[x,y]$ for the line segment $\{\lambda x + (1-\lambda)y \colon 0 \leq \lambda \leq 1\}$ between x and y. Then the set

$$X := [(0,0),(0,1)] \cup \bigcup_{n\in\mathbb{N}} [(\tfrac{1}{n},0),(0,1)]$$

is arcwise connected but not locally connected (in the points $(0,x), 0 \leq x < 1$).

(c) The set

$$X := (\{0\} \times [-1,1]) \cup \{(x, \sin\tfrac{1}{x} \colon 0 < x < 1\} \subseteq \mathbb{R}^2$$

is connected but not arcwise connected.

Exercises for Appendix A

Exercise A.5.1. (a) For every system $\mathcal{A}$ of subsets of a set X with $\bigcup \mathcal{A} = X$ the system $\widetilde{\mathcal{A}}$ of all finite intersections of sets in $\mathcal{A}$ is a basis for the topology generated by $\mathcal{A}$.

(b) A system $\mathcal{A}$ of subsets of X is the basis of a topology if and only if

(B1) $\bigcup \mathcal{A} = X$ and

(B2) for each $x \in A \cap B$, $A, B \in \mathcal{A}$, there exists a $C \in \mathcal{A}$ with $x \in C \subseteq A \cap B$.

Exercise A.5.2. (a) If $f \colon X \to Z$ is a continuous map and $Y \subseteq X$ a subset endowed with the subspace topology, then the restriction $f|_Y \colon Y \to Z$ is continuous.

(b) Let $X_1, \ldots, X_n$ be topological spaces and $A_j \subseteq X_j$ subsets. Show that

$$\overline{A_1 \times \ldots \times A_n} = \overline{A_1} \times \ldots \times \overline{A_n}$$

holds in $X := X_1 \times \ldots \times X_n$ with respect to the product topology and likewise

$$(A_1 \times \ldots \times A_n)^0 = A_1^0 \times \ldots \times A_n^0.$$

(c) Let X and Y be topological spaces. Then for each $x \in X$ the map

$$j_x \colon Y \to X \times Y, \quad y \mapsto (x,y)$$

is continuous and the corestriction

$$j_x^{|Y\times\{x\}} \colon Y \to Y \times \{x\}$$

is a homeomorphism.

Exercise A.5.3. (a) If (X_i, d_i), $i = 1, \ldots, n$, are metric spaces, then the product topology on $X = X_1 \times \ldots \times X_n$ is induced by the following metrics:

(1) $d(x,y) := \sum_{i=1}^n d_i(x_i, y_i)$.

(2) $d(x,y) := \max(d_1(x_1,y_1), \ldots, d_n(x_n,y_n))$.

(b) A sequence $(x^{(m)})_{m\in\mathbb{N}}$ in X converges to $x = (x_1, \ldots, x_n)$ if and only if all component sequences $(x_j^{(m)})_{m\in\mathbb{N}}$ converge to x_j.

Exercise A.5.4. Let X and Y be topological spaces and $X \times Y$ their topological product. Then the connected components and the arc components in $X \times Y$ are given by

$$C(x,y) = C(x) \times C(y) \quad \text{and} \quad C_a(x,y) = C_a(x) \times C_a(y)$$

for $(x,y) \in X \times Y$. In particular, the product space $X \times Y$ is connected, resp., arcwise connected if and only if X and Y are connected, resp., arcwise connected.

Exercise A.5.5. In $\mathbb{R}^2$ we consider the set

$$X = ([0,1] \times \{1\}) \cup \left(\left\{\frac{1}{n} : n \in \mathbb{N}\right\} \times [0,1]\right) \cup (\{0\} \times [0,1]).$$

Show that X is arcwise connected but not locally arcwise connected.

Appendix B

Analytic functions

In this appendix we recall the concept of differentiability, formulated in the context of Banach spaces. In particular, we discuss the concept of an analytic function, which provides a very direct tool to verify the smoothness of functions given by non-commutative power series, such as the exponential function of a Banach algebra.

B.1 Differentiable functions

We briefly recall the concept of a differentiable function between open subsets of Banach spaces. For two normed spaces we write $\mathcal{L}(X, Y)$ for the vector space of continuous linear mas $A\colon X \to Y$ endowed with the operator norm

$$\|A\| := \sup\{\|Ax\| \colon x \in X, \|x\| \leq 1\}.$$

Definition B.1.1. Let X and Y be Banach spaces and $U \subseteq X$ an open subset. We say that a map $f\colon U \to X$ is *differentiable in* $x \in U$ if there exists a continuous linear map $\mathtt{d}f(x) \in \mathcal{L}(X, Y)$ with

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - \mathtt{d}f(x)(h)\|}{\|h\|} = 0. \qquad \text{(B.1)}$$

This implies in particular, for each $v \in X$, the relation

$$\mathtt{d}f(x)(v) = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$

which shows that we can interprete the value $\mathtt{d}f(x)(v)$ as the derivative of f in x in the direction of v, and moreover, that $\mathtt{d}f(x)$ is uniquely determined as a continuous linear map satisfying (2.1): Whenever there exists a continuous linear map $A \in \mathcal{L}(X, Y)$ with

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - A(h)\|}{\|h\|} = 0,$$

then f is differentiable in x and $\mathtt{d}f(x) = A$.

We call f *differentiable* if f is differentiable in every point $x \in U$. Then we obtain a map $\mathtt{d}f\colon U \to \mathcal{L}(X,Y)$, and we call f a C^1-map if $\mathtt{d}f$ is a continuous map. By iteration, we define C^k-maps for $k \in \mathbb{N}$ as C^1-maps for which $\mathtt{d}f$ is C^{k-1}. We call a map *smooth* or a C^∞-map if it is C^k for every $k \in \mathbb{N}$.

Remark B.1.2. (a) If $f\colon U \to Y$ is differentiable in $x \in U$, then (B.1) implies that there exists a $\delta > 0$ with $B_\delta(x) \subseteq U$ and

$$\|f(x+h) - f(x) - \mathtt{d}f(x)(h)\| \leq \|h\|$$

for $\|h\| < \delta$. This implies that

$$\|f(x+h) - f(x)\| \leq \|\mathtt{d}f(x)\|\|h\| + \|h\|,$$

and therefore f is continuous in x. Hence differentiable functions are continuous.

(b) If $f\colon X \to Y$ is a continuous linear map, then f is differentiable with $\mathtt{d}f(x) = f$ for each $x \in X$. We therefore observe that the differential of a linear map is constant. If f is constant, then its differential obviously vanishes. In the next subsection we shall study bilinear maps, which have the property that their differential, viewed as a map $\mathtt{d}f\colon U \to \mathcal{L}(X,Y)$ is linear.

Theorem B.1.3. (Chain Rule) *Let X, Y, Z be Banach spaces, $U \subseteq X$ and $V \subseteq Z$ open and $f\colon V \to Z$, $g\colon U \to V$ maps such that g is differentiable in x and f is differentiable in $g(x)$. Then $f \circ g$ is differentiable in x with*

$$\mathtt{d}(f \circ g)(x) = \mathtt{d}f(g(x)) \circ \mathtt{d}g(x).$$

Proof. For $y := g(x)$ we write

$$g(x+h) = g(x) = \mathtt{d}g(x)h + r_g(h) \quad \text{and} \quad f(y+h) = f(y) = \mathtt{d}f(y)h + r_f(h)$$

with $\frac{\|r_f(h)\|}{\|h\|} \to 0$ and $\frac{\|r_g(h)\|}{\|h\|} \to 0$ for $h \to 0$. For $x + h \in U$ we then obtain

$$\begin{aligned} f(g(x+h)) - f(g(x)) &= f(y + \mathtt{d}g(x)h + r_g(h)) - f(y) \\ &= \mathtt{d}f(y)\mathtt{d}g(x)h + \mathtt{d}f(y)r_g(h) + r_f(\mathtt{d}g(x)h + r_g(h)). \end{aligned}$$

Since $\mathtt{d}f(y)$ is a continuous linear map, we obtain

$$\frac{\|\mathtt{d}f(y)r_g(h)\|}{\|h\|} \leq \|\mathtt{d}f(y)\|\frac{\|r_g(h)\|}{\|h\|} \to 0$$

for $h \to 0$. To see that

$$\lim_{h \to 0} \frac{\|r_f(\mathtt{d}g(x)h + r_g(h))\|}{\|h\|} = 0, \tag{B.2}$$

let $\varepsilon > 0$. Then there exists a $\delta > 0$ such that $\|r_f(z)\| < \varepsilon\|z\|$ for $\|z\| \leq \delta$. Since $\lim_{h \to 0} \mathtt{d}g(x)h + r_g(h) = 0$, there exists an $\eta > 0$ such that the norm of this expression is $< \delta$ for $\|h\| < \eta$. For $\|h\| < \eta$ we then have

$$\frac{\|r_f(\mathtt{d}g(x)h + r_g(h))\|}{\|h\|} \leq \varepsilon\frac{\|\mathtt{d}g(x)h + r_g(h)\|}{\|h\|} \leq \varepsilon\|\mathtt{d}g(x)\| + \varepsilon\frac{\|r_g(h)\|}{\|h\|}.$$

This implies (B.2). □

B.2 Multilinear maps

Definition B.2.1. Let $X_1, \ldots, X_n$ and Y be normed spaces. A function

$$\beta \colon X_1 \times \cdots \times X_n \to Y$$

is called *multilinear* or *n-linear* if for each j and fixed elements $x_i \in X_i$ for $i \neq j$ the map

$$X_j \to Y, \quad x \mapsto \beta(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n)$$

is linear.

We define the norm of a multilinear map by

$$\|\beta\| := \sup\{\|\beta(x_1, \ldots, x_n)\| \colon \|x_j\| \leq 1, j = 1, \ldots, n\} \in [0, \infty].$$

Lemma B.2.2. *For a multilinear map $\beta \colon X_1 \times \cdots \times X_n \to Y$ the following are equivalent:*

(i) *β is continuous.*

(ii) *β is continuous in $(0, \ldots, 0)$.*

(iii) $\|\beta\| < \infty$.

If these conditions are satisfied, then we have

$$\|\beta(x_1, \ldots, x_n)\| \leq \|\beta\| \cdot \|x_1\| \cdots \|x_n\| \quad \textit{for} \quad x_j \in X_j, \ j = 1, \ldots, n.$$

Proof. (i) $\Rightarrow$ (ii) is trivial.

(ii) $\Rightarrow$ (iii): The continuity of β in $(0, \ldots, 0)$ implies in particular that there exists a neighborhood U of $(0, \ldots, 0)$ in the product space $X_1 \times \ldots \times X_n$ such that

$$\|\beta(x_1, \ldots, x_n)\| \leq 1$$

for $(x_1, \ldots, x_n) \in U$. On the other hand there exists a $\delta > 0$ with

$$(\forall j)\, \|x_j\| \leq \delta \quad \Rightarrow \quad (x_1, \ldots, x_n) \in U.$$

This implies that $\|\beta\| \leq \delta^{-n}$.

(iii) $\Rightarrow$ (i): In view of

$$\begin{aligned}
&\beta(x_1, x_2, \ldots, x_n) - \beta(x_1', x_2', \ldots, x_n') \\
=&\big(\beta(x_1, x_2, \ldots, x_n) - \beta(x_1', x_2, \ldots, x_n)\big) \\
&\quad + \big(\beta(x_1', x_2, \ldots, x_n) - \beta(x_1', x_2', \ldots, x_n)\big) \\
&\quad + \ldots + \big(\beta(x_1', x_2', \ldots, x_{n-1}', x_n) - \beta(x_1', x_2', \ldots, x_{n-1}', x_n')\big) \\
=&\beta(x_1 - x_1', x_2, \ldots, x_n) + \beta(x_1', x_2 - x_2', \ldots, x_n) + \cdots + \beta(x_1', x_2', \ldots, x_n - x_n') \\
=&\sum_{j=1}^{n} \beta(x_1', \ldots, x_{j-1}', x_j - x_j', x_{j+1}, \ldots, x_n),
\end{aligned}$$

we have

$$\|\beta(x) - \beta(x')\| \le \|\beta\| \sum_{j=1}^{n} \|x_1\| \cdots \|x_{j-1}\| \|x_j - x'_j\| \|x'_{j+1}\| \cdots \|x'_n\|.$$

This implies the continuity of β. □

One easily verifies that $\|\cdot\|$ defines a norm on the space $\mathcal{L}(X_1, \ldots, X_n; Y)$ of continuous multilinear maps from $X_1 \times \ldots \times X_n$ to Y. If $X_1 = \ldots = X_n = X$, then we write $\mathcal{L}^n(X; Y) := \mathcal{L}(X, \ldots, X; Y)$ for the space of continuous n-linear maps $X^n \to Y$. For $n = 0$ we put $\mathcal{L}^0(X; Y) := Y$ and note that $\mathcal{L}^1(X; Y) = \mathcal{L}(X, Y)$ is the space of continuous linear maps, endowed with the operator norm.

We conclude this subsection with a discussion of the differentiability properties of multilinear maps. The following lemma gives the formula for the derivative, which is an important tool to calculate derivatives of complicated matrix-valued or operator-valued maps.

Lemma B.2.3. *Each continuous multilinear map $\beta\colon X_1 \times \cdots \times X_n \to Y$ is differentiable with*

$$\mathtt{d}\beta(x_1, \ldots, x_n)(h_1, \ldots, h_n) = \beta(h_1, x_2, \ldots, x_n) + \ldots + \beta(x_1, x_2, \ldots, h_n). \tag{B.3}$$

Proof. Fix $x = (x_1, \ldots, x_n)$ and define $A := \mathtt{d}\beta(x_1, \ldots, x_n)$ by (B.3). On the product space $X := X_1 \times \ldots \times X_n$ we consider the norm $\|x\| := \max_j \|x_j\|$. Then $A\colon X \to Y$ is a continuous linear map with

$$\|A\| \le \|\beta\|(\|x_2\| \cdots \|x_n\| + \cdots + \|x_1\| \cdots \|x_{n-1}\|) \le n\|\beta\|\|x\|^{n-1}.$$

Moreover, the additive expansion of $\beta(x + h)$ yields for $j = 2, \ldots, n$ summands $\gamma_j h^j$ where each γ_j is j-linear with

$$\|\gamma_j\| \le \binom{n}{j} \|\beta\|\|x\|^{n-j}.$$

Therefore

$$\|\beta(x + h) - \beta(x) - A(h)\| \le \sum_{j=2}^{n} \binom{n}{j} \|\beta\|\|x\|^{n-j}\|h\|^j,$$

which implies that

$$\lim_{\|h\| \to 0} \frac{\|\beta(x + h) - \beta(x) - A(h)\|}{\|h\|} = 0.$$ □

Example B.2.4. (a) If $\beta\colon X_1 \times X_2 \to Y$ is a continuous bilinear map, then we have

$$\mathtt{d}\beta(x_1, x_2)(h_1, h_2) = \beta(h_1, x_2) + \beta(x_1, h_2).$$

In particular, we observe that for $X := X_1 \times X_2$ the map

$$\mathtt{d}\beta\colon X \to \mathcal{L}(X, Y)$$

is a continuous linear map.

Now let $\beta\colon X \times X \to Y$ be a continuous bilinear map and consider the corresponding quadratic map $q\colon X \to Y, q(x) := \beta(x,x)$. Then we write $q = \beta \circ \Delta$, where $\Delta\colon X \to X \times X$ is the diagonal embedding. Since Δ is linear, its differential is simply given by

$$\mathtt{d}\Delta(x)(h) = \Delta(h) = (h,h).$$

Therefore the chain rule leads to

$$\mathtt{d}q(x)(h) = \mathtt{d}\beta(\Delta(x))\mathtt{d}\Delta(x)(h) = \mathtt{d}\beta(x,x)(h,h) = \beta(h,x) + \beta(x,h).$$

If, in addition, β is symmetric, then this formula simplifies to

$$\mathtt{d}q(x)(h) = 2\beta(x,h).$$

(b) If $\beta\colon X^n \to Y$ is a continuous multilinear map and $q(x) := \beta(x,\ldots,x)$, then similar arguments as in (a) lead to

$$\mathtt{d}q(x)(h) = \mathtt{d}\beta(x,\ldots,x)(h,\ldots,h) = \beta(h,x,\ldots,x) + \cdots + \beta(x,\ldots,x,h).$$

We call β *symmetric* if for every permutation $\sigma \in S_n$ we have

$$\beta(x_{\sigma(1)},\ldots,x_{\sigma(n)}) = \beta(x_1,\ldots,x_n), \quad x_1,\ldots,x_n \in X.$$

If β is symmetric, then we immediately get the simpler formula

$$\mathtt{d}q(x)(h) = n\beta(x,\ldots,x,h).$$

(c) If $(\mathcal{A}, \|\cdot\|)$ is a Banach algebra, then we consider the power maps

$$p_n\colon \mathcal{A} \to \mathcal{A}, \quad p_n(a) = a^n.$$

Since the n-fold multiplication map

$$\beta_n\colon \mathcal{A}^n \to \mathcal{A}, \quad (a_1,\ldots,a_n) \mapsto a_1 \cdots a_n$$

is multilinear and continuous with $\|\beta_n\| \leq 1$, we can use (b) to calculate the derivative of p_n as

$$\mathtt{d}p_n(a)(h) = \beta(h,a,\ldots,a) + \cdots + \beta(a,\ldots,a,h) = ha^{n-1} + aha^{n-2} + \cdots + a^{n-1}h.$$

If, in addition, the multiplication is commutative, then we obtain the simpler formula

$$\mathtt{d}p_n(a)(h) = na^{n-1}h$$

which more reassembles the formula one learns in calculus courses for the derivative of the power functions on the algebras $\mathcal{A} = \mathbb{R}, \mathbb{C}$.

B.3 Analytic functions

Definition B.3.1. Let X and Y be Banach spaces and $U \subseteq X$ an open subset. For $c_n \in \mathcal{L}^n(X;Y)$ and $x \in X$ we define $c_n x^n := c_n(x, \dots, x)$. A map $f\colon U \to Y$ is called *analytic* if for each point $x_0 \in X$ there exist $c_n \in \mathcal{L}^n(X;Y)$ such that for some $r > 0$ we have

(A1) $\sum_{n=1}^{\infty} \|c_n\| r^n < \infty$, and

(A2) $f(x_0 + h) = f(x_0) + \sum_{n=1}^{\infty} c_n h^n$ for $\|h\| < r$ and $x_0 + h \in U$.

Remark B.3.2. (a) For $X = \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and Y a $\mathbb{K}$-Banach space each multilinear map $c_n\colon X^n \to Y$ satisfies

$$c_n(z_1, \dots, z_n) = z_1 \cdots z_n \cdot c_n(1, \dots, 1),$$

so that $c_n(z, \dots, z) = z^n c_n(1, \dots, 1)$ and $\|c_n\| = |c_n(1, \dots, 1)|$. Therefore analytic functions are those which can be represented locally by power series of the type $\sum_{n=0}^{n} z^n a_n$ with $a_n \in Y$.

(b) For $X = \mathbb{K}^m$ an n-linear map $c_n\colon X^n \to Y$ is a sum of terms of the type

$$z_1, \dots, z_n \mapsto z_{1,j_1} \dots z_{n,j_n} a_{j_1,\dots,j_n},$$

where $a_{j_1,\dots,j_n} \in Y$ and $j_1, \dots, j_n \in \{1, \dots, m\}$ denote the components of $z_j \in \mathbb{K}^m$. We therefore find expressions which are familiar from the Taylor expansion in several variables.

Lemma B.3.3. *Analytic functions $f\colon U \to Y$ are differentiable in each point of U.*

Proof. Let $f\colon U \to Y$ be analytic and $x_0 \in U$. For $\|h\| < r$ as (A2) above, we then obtain

$$\|f(x_0 + h) - f(x_0)\| \le \sum_{n=1}^{\infty} \|c_n\| \|h\|^n = \|h\| \sum_{n=0}^{\infty} \|c_{n+1}\| \|h\|^n,$$

and therefore the continuity of f in x_0. Moreover, we have

$$\|f(x_0 + h) - f(x_0) - c_1(h)\| \le \sum_{n=2}^{\infty} \|c_n\| \|h\|^n = \|h\|^2 \sum_{n=0}^{\infty} \|c_{n+2}\| \|h\|^n,$$

so that

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - c_1(h)\|}{\|h\|} = 0.$$

This implies the differentiability of f in x_0 with $\mathtt{d}f(x_0) = c_1$. $\square$

Now the natural question is whether the derivatives of analytic functions are again analytic, and furthermore, how we can see whether a concretely given function is analytic or not. Obviously it would not be reasonable to verify the condition in Definition B.3.1 in every point of U. The following proposition is the crucial observation.

Proposition B.3.4. *Let X and Y be Banach spaces and $c_n \in \mathcal{L}^n(X;Y)$, $n \in \mathbb{N}_0$, with $\sum_{n=0}^{\infty} \|c_n\| r^n < \infty$. Then*

$$f\colon B_r(0) \to Y, \quad f(x) := \sum_{n=0}^{\infty} c_n x^n$$

defines an analytic function.

Proof. By expanding multilinearly, we get for $n = 3$:

$$\begin{aligned} c_3(x+y)^3 = c_3(x,x,x) &+ \big(c_3(y,x,x) + c_3(x,y,x) + c_3(x,x,y)\big) \\ &+ \big(c_3(y,y,x) + c_3(y,x,y) + c_3(x,y,y)\big) + c_3(y,y,y), \end{aligned}$$

and more generally

$$c_n(x+y)^n = d_{n,0}(x) + d_{n,1}(x)y + d_{n,2}(x)y^2 + \ldots + d_{n,n}(x)y^n,$$

with $d_{n,j}(x) \in \mathcal{L}^j(X;Y)$ and

$$\|d_{n,j}(x)\| \leq \|c_n\| \|x\|^{n-j} \binom{n}{j}.$$

Suppose that $\|x\| < r$ and $s < r - \|x\|$. Then we have

$$\sum_{j=0}^{n} \|x\|^{n-j} \binom{n}{j} s^j = (\|x\| + s)^n < r^n,$$

and therefore

$$\sum_{j \leq n} \|c_n\| \|x\|^{n-j} \binom{n}{j} s^j \leq \sum_n \|c_n\| r^n < \infty.$$

We conclude in particular that for $\|x\| < r$ the series

$$c_j^x := \sum_{n \geq j} d_{n,j}(x)$$

converges in $\mathcal{L}^j(X;Y)$ with

$$\|c_j^x\| \leq \sum_{n \geq j} \|d_{n,j}(x)\| \leq \sum_{n \geq j} \binom{n}{j} \|c_n\| \|x\|^{n-j},$$

so that

$$\sum_j \|c_j^x\| s^j \leq \sum_{n \geq j} \binom{n}{j} \|c_n\| \|x\|^{n-j} s^j < \infty.$$

This implies that for $\|y\| < s$ the series

$$h(y) := \sum_{j=0}^{\infty} c_j^x y^n$$

converges. On the other hand, absolute convergence of the series under consideration implies that we may rearrange the summation to obtain

$$f(x+y) = \sum_n c_n(x+y)^n = \sum_n \sum_{j \leq n} d^x_{n,j} y^j = \sum_j \big(\sum_{n \geq j} d^x_{n,j}\big) y^j = \sum_j c^x_j y^j = h(y)$$

(Exercise 2.1.1) Therefore f can be represented on $B_s(x)$ by a power series as in (A2), and therefore f is analytic. □

Remark B.3.5. The proof of Proposition B.3.4 further shows that for each $x \in B_r(0)$ we have

$$\mathtt{d}f(x)(y) = c^x_1(y) = \sum_{n=1}^{\infty} d_{n,1}(x)(y).$$

Viewing $d_{n,1}$ as an element in $\mathcal{L}^{n-1}(X; \mathcal{L}(X,Y))$, we have

$$\|d_{n,1}\| \leq \|c_n\| n,$$

and therefore

$$\sum_n \|d_{n,1}\| r^n < \infty.$$

This implies that the derivative $\mathtt{d}f \colon U \to \mathcal{L}(X,Y)$ of an analytic function on U is also analytic and that its local power series expansion can be obtained from the expansion of f by taking the derivative c_n^1 of each summand c_n.

This discussion leads immediately to the following corollary:

Corollary B.3.6. *Analytic functions are smooth.*

Combining Proposition B.3.4 with Corollary B.3.6, we obtain:

Theorem B.3.7. *Let X and Y be Banach spaces and $c_n \in \mathcal{L}^n(X;Y)$, $n \in \mathbb{N}_0$, with $\sum_{n=0}^{\infty} \|c_n\| r^n < \infty$. Then*

$$f \colon B_r(0) \to Y, \quad f(x) := \sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} c_n(x, \ldots, x)$$

defines a smooth function whose derivative is given by

$$\mathtt{d}f(x) = \sum_{n=0}^{\infty} \mathtt{d}c_n(x).$$

The following theorem is a central result on analytic functions. It is false for smooth functions, and in this sense it describes a property which is characteristic for analytic functions.

Theorem B.3.8. (Identity Theorem for Analytic Functions) *Let X and Y be Banach spaces, $U \subseteq X$ open and connected, and $f, g \colon U \to Y$ analytic functions. If $f = g$ holds on an open subset $V \subseteq U$, then $f = g$ holds on U.*

Proof. Replacing f by $f - g$, we may assume that $g = 0$. We consider the open set $V := f^{-1}(0)^0$. Then our assumption implies that $V \neq \emptyset$.

We claim that V is closed. So let $x_0 \in \overline{V}$, and suppose that $B_{2r}(x_0) \subseteq U$ and that for $\|h\| < 2r$ we have the power series expansion $f(x_0 + h) = \sum_{n=0}^\infty c_n h^n$. As in the proof of Proposition B.3.4, we then obtain for for $\|x - x_0\|$ and $\|h\| < r$ the expansion

$$f(x+h) = \sum_{j=0}^\infty c_j^x h^j,$$

where the functions

$$c_j \colon B_r(x_0) \to \mathcal{L}^j(X;Y), \quad x \mapsto c_j^x$$

are analytic, hence in particular continuous, with $c_j^{x_0} = c_j$.

For $|t| < 1$ we have

$$f(x+th) = \sum_{j=0}^\infty c_j^x(h,\ldots,h)t^j.$$

In view of $x \in V$, this real analytic function vanishes in a neighborhood of 0. Hence all its derivatives vanish, and we obtain

$$c_j^x(h,\ldots,h) = \frac{1}{j!}\frac{d^n f(x+th)}{dt^n}(0) = 0.$$

Since x_0 is the limit of a sequence $x_n \in V$, we obtain

$$c_j^{x_0}(h,\ldots,h) = 0,$$

and therefore

$$f(x_0 + h) = \sum_j c_j^{x_0} h^j = 0$$

for $\|h\| < r$. This means that $x_0 \in V$, and hence that V is closed.

Therefore V is a non-empty open closed subset of U, hence coincides with U because U is connected, and this means that $f = 0$ on U. □

For the applications of analytic functions it is important that the composition of two analytic functions is again analytic.

Theorem B.3.9. *Let X, Y, Z be Banach spaces, $U \subseteq X$ and $V \subseteq Z$ open and $g \colon V \to Z$, $f \colon U \to V$ analytic maps. Then the map $g \circ f \colon U \to Z$ is analytic.*

Proof. Let $x \in U$, $y := f(x) \in V$ and $r > 0$ such that

$$f(x+h) = \sum_n c_n h^n \quad \text{and} \quad g(y+h) = \sum_n d_n h^n$$

for $\|h\| < r$ with $\sum_n \|c_n\| r^n < \infty$ and $\sum_n \|d_n\| r^n < \infty$.

First we consider the compositions $d_m \circ (f - f(x)) \colon U \to Z$. Since d_m is a continuous m-linear function, $d_m(\sum_{n=1}^{\infty} c_n h^n)$ is defined for $\|h\| < r$. Moreover, it can be written as a series

$$d_m\Big(\sum_{n=1}^{\infty} c_n h^n\Big) = \sum_{j=m}^{\infty} d_m^j h^j,$$

where $d_m^m = d_m \circ c_1$ and the other terms are obtained by collecting all terms of the same degree in h. We thus obtain $d_m^j \in \mathcal{L}^j(X; Z)$ with

$$\|d_m^j\| \leq \|d_m\| \sum_{j_1+\ldots+j_m=j} \|c_{j_1}\| \cdots \|c_{j_m}\|.$$

For $s > 0$ this further leads to

$$\begin{aligned}\sum_{j=m}^{\infty} \|d_m^j\| s^j &\leq \|d_m\| \sum_{j_1+\ldots+j_m \geq m} \|c_{j_1}\| \cdots \|c_{j_m}\| s^{j_1+\ldots+j_m} \\ &\leq \|d_m\| \sum_{j_1+\ldots+j_m \geq m} \|c_{j_1}\| s^{j_1} \cdots \|c_{j_m}\| s^{j_m} \\ &\leq \|d_m\| \Big(\sum_{j=1}^{\infty} \|c_j\| s^j\Big)^m.\end{aligned}$$

Summing also over m, we eventually get

$$\sum_m \sum_{j=m}^{\infty} \|d_m^j\| s^j \leq \sum_m \|d_m\| \Big(\sum_{j=1}^{\infty} \|c_j\| s^j\Big)^m < \infty$$

for $\sum_{j=1}^{\infty} \|c_j\| s^j < r$, which is the case if $s > 0$ is small enough. Therefore

$$f(g(x+h)) = \sum_n e_n h^n \quad \text{with} \quad e_n := \sum_{m \leq n} d_m^n \quad \text{and} \quad \sum_n \|e_n\| s^n < \infty.$$

This proves that $f \circ g$ is an analytic function. □

Exercises for Appendix B

Exercise B.3.1. If $X_1, \ldots, X_n$ are finite-dimensional normed spaces, then each multilinear map $\beta \colon X_1 \times \ldots \times X_n \to Y$ is continuous. Hint: Choose a basis in each space X_j and expand β accordingly.

Appendix C

Covering Theory

In this appendix we provide the main results on coverings of topological spaces needed in particular to calculate fundamental groups and to prove the existence of simply connected covering spaces.

C.1 The Fundamental Group

To define the notion of a simply connected space, we first have to define its fundamental group. The elements of this group are homotopy classes of loops. The present section develops this concept and provides some of its basic properties.

Definition C.1.1. Let X be a topological space, $I := [0, 1]$, and $x_0, x_1 \in X$. We write

$$P(X, x_0) := \{\gamma \in C(I, X) \colon \gamma(0) = x_0\}$$

and

$$P(X, x_0, x_1) := \{\gamma \in P(X, x_0) \colon \gamma(1) = x_1\}.$$

We call two paths $\alpha_0, \alpha_1 \in P(X, x_0, x_1)$ *homotopic*, written $\alpha_0 \sim \alpha_1$, if there exists a continuous map

$$H \colon I \times I \to X \quad \text{with} \quad H_0 = \alpha_0, \quad H_1 = \alpha_1$$

(for $H_t(s) := H(t, s)$) and

$$(\forall t \in I) \quad H(t, 0) = x_0, \ H(t, 1) = x_1.$$

It is easy to show that $\sim$ is an equivalence relation (Exercise C.1.2), called *homotopy*. The homotopy class of α is denoted by $[\alpha]$.

We write $\Omega(X, x_0) := P(X, x_0, x_0)$, for the set of loops based at x_0. For $\alpha \in P(X, x_0, x_1)$ and $\beta \in P(X, x_1, x_2)$ we define a product $\alpha * \beta$ in $P(X, x_0, x_2)$ by

$$(\alpha * \beta)(t) := \begin{cases} \alpha(2t) & \text{for } 0 \leq t \leq \frac{1}{2} \\ \beta(2t - 1) & \text{for } \frac{1}{2} \leq t \leq 1. \end{cases}$$

Lemma C.1.2. *If $\varphi\colon [0,1] \to [0,1]$ is a continuous map with $\varphi(0) = 0$ and $\varphi(1) = 1$, then for each $\alpha \in P(X, x_0, x_1)$ we have $\alpha \sim \alpha \circ \varphi$.*

Proof. Use $H(t, s) := \alpha(ts + (1-t)\varphi(s))$. □

Proposition C.1.3. *The following assertions hold:*

(1) *$\alpha_1 \sim \alpha_2$ and $\beta_1 \sim \beta_2$ implies $\alpha_1 * \beta_1 \sim \alpha_2 * \beta_2$, so that we obtain a well-defined product*

$$[\alpha] * [\beta] := [\alpha * \beta]$$

of homotopy classes.

(2) *If x also denotes the constant map $I \to \{x\} \subseteq X$, then*

$$[x_0] * [\alpha] = [\alpha] = [\alpha] * [x_1] \quad \text{for} \quad \alpha \in P(X, x_0, x_1).$$

(3) *(Associativity) $[\alpha * \beta] * [\gamma] = [\alpha] * [\beta * \gamma]$ for $\alpha \in P(X, x_0, x_1)$, $\beta \in P(X, x_1, x_2)$ and $\gamma \in P(X, x_2, x_3)$.*

(4) *(Inverse) For $\alpha \in P(X, x_0, x_1)$ and $\overline{\alpha}(t) := \alpha(1-t)$ we have*

$$[\alpha] * [\overline{\alpha}] = [x_0].$$

(5) *(Functoriality) For any continuous map $\varphi\colon X \to Y$ and $\alpha \in P(X, x_0, x_1), \beta \in P(X, x_1, x_2)$, we have*

$$(\varphi \circ \alpha) * (\varphi \circ \beta) = \varphi \circ (\alpha * \beta),$$

and $\alpha \sim \beta$ implies $\varphi \circ \alpha \sim \varphi \circ \beta$.

Proof. (1) If H^α is a homotopy from α_1 to α_2 and H^β a homotopy from β_1 to β_2, then we put

$$H(t, s) := \begin{cases} H^\alpha(t, 2s) & \text{for } 0 \leq s \leq \frac{1}{2} \\ H^\beta(t, 2s-1) & \text{for } \frac{1}{2} \leq s \leq 1 \end{cases}$$

(cf. Exercise C.1.1).

(2) For the first assertion we use Lemma C.1.2 and

$$x_0 * \alpha = \alpha \circ \varphi \quad \text{for} \quad \varphi(t) := \begin{cases} 0 & \text{for } 0 \leq t \leq \frac{1}{2} \\ 2t - 1 & \text{for } \frac{1}{2} \leq t \leq 1. \end{cases}$$

For the second, we have

$$\alpha * x_1 = \alpha \circ \varphi \quad \text{for} \quad \varphi(t) := \begin{cases} 2t & \text{for } 0 \leq t \leq \frac{1}{2} \\ 1 & \text{for } \frac{1}{2} \leq t \leq 1. \end{cases}$$

(3) We have $(\alpha * \beta) * \gamma = (\alpha * (\beta * \gamma)) \circ \varphi$ for

$$\varphi(t) := \begin{cases} 2t & \text{for } 0 \leq t \leq \frac{1}{4} \\ \frac{1}{4} + t & \text{for } \frac{1}{4} \leq t \leq \frac{1}{2} \\ \frac{t+1}{2} & \text{for } \frac{1}{2} \leq t \leq 1. \end{cases}$$

(4)
$$H(t,s) := \begin{cases} \alpha(2s) & \text{for } s \le \frac{1-t}{2} \\ \alpha(1-t) & \text{for } \frac{1-t}{2} \le s \le \frac{1+t}{2} \\ \overline{\alpha}(2s-1) & \text{for } s \ge \frac{1+t}{2}. \end{cases}$$

(5) is trivial. □

Definition C.1.4. From the preceding definition, we derive in particular that the set

$$\pi_1(X, x_0) := \Omega(X, x_0)/\sim$$

of homotopy classes of loops in x_0 carries a natural group structure, given by

$$[\alpha][\beta] := [\alpha * \beta]$$

(Exercise). This group is called the *fundamental group of X with respect to x_0*.

A pathwise connected space X is called *simply connected* if $\pi_1(X, x_0)$ vanishes for some $x_0 \in X$ (which implies that is trivial for each $x_0 \in X$; Exercise C.1.4).

Lemma C.1.5. (Functoriality of the fundamental group) *If $f\colon X \to Y$ is a continuous map with $f(x_0) = y_0$, then*

$$\pi_1(f, x_0)\colon \pi_1(X, x_0) \to \pi_1(Y, y_0), \quad [\gamma] \mapsto [f \circ \gamma]$$

is a group homomorphism. Moreover, we have

$$\pi_1(\mathrm{id}_X, x_0) = \mathrm{id}_{\pi_1(X,x_0)} \quad \text{and} \quad \pi_1(f \circ g, x_0) = \pi_1(f, g(x_0)) \circ \pi_1(g, x_0).$$

Proof. This follows directly from Proposition C.1.3(5). □

Remark C.1.6. The map

$$\sigma\colon \pi_1(X, x_0) \times \big(P(X, x_0)/\sim\big) \to P(X, x_0)/\sim, \quad ([\alpha], [\beta]) \mapsto [\alpha * \beta] = [\alpha] * [\beta]$$

defines an action of the group $\pi_1(X, x_0)$ on the set $P(X, x_0)/\sim$ of homotopy classes of paths starting in x_0 (Proposition C.1.3).

Remark C.1.7. (a) Suppose that the topological space X is contractible, i.e., there exists a continuous map $H\colon I \times X \to X$ and $x_0 \in X$ with $H(0, x) = x$ and $H(1, x) = x_0$ for $x \in X$. Then $\pi_1(X, x_0) = \{[x_0]\}$ is trivial (Exercise).
(b) $\pi_1(X \times Y, (x_0, y_0)) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0)$ (Exercise).
(c) $\pi_1(\mathbb{R}^n, 0) = \{0\}$ because $\mathbb{R}^n$ is contractible.

More generally, if the open subset $\Omega \subseteq E$ of the Banach space E is starlike with respect to x_0, then $H(t, x) := x + t(x - x_0)$ yields a contraction to x_0, and we conclude that $\pi_1(\Omega, x_0)$ is trivial.

The following lemma implies in particular, that fundamental groups of topological groups are always abelian.

Lemma C.1.8. *Let G be a topological group and consider the identity element $\mathbf{1}$ as a base point. Then the path space $P(G, \mathbf{1})$ also carries a natural group structure given by the pointwise product $(\alpha \cdot \beta)(t) := \alpha(t)\beta(t)$ and we have*

(1) *$\alpha \sim \alpha'$, $\beta \sim \beta'$ implies $\alpha \cdot \beta \sim \alpha' \cdot \beta'$, so that we obtain a well-defined product*

$$[\alpha] \cdot [\beta] := [\alpha \cdot \beta]$$

of homotopy classes, defining a group structure on $P(G, \mathbf{1})/\sim$.

(2) *$\alpha \sim \beta \Longleftrightarrow \alpha \cdot \beta^{-1} \sim \mathbf{1}$, the constant map.*

(3) *(Commutativity) $[\alpha] \cdot [\beta] = [\beta] \cdot [\alpha]$ for $\alpha, \beta \in \Omega(G, \mathbf{1})$.*

(4) *(Consistency) $[\alpha] \cdot [\beta] = [\alpha] * [\beta]$ for $\alpha \in \Omega(G, \mathbf{1})$, $\beta \in P(G, \mathbf{1})$.*

Proof. (1) follows by composing homotopies with the multiplication map m_G.
(2) follows from (1) by multiplication with β^{-1}.
(3)

$$[\alpha][\beta] = [\alpha * \mathbf{1}][\mathbf{1} * \beta] = [(\alpha * \mathbf{1})(\mathbf{1} * \beta)] = [(\mathbf{1} * \beta)(\alpha * \mathbf{1})] = [\mathbf{1} * \beta][\alpha * \mathbf{1}] = [\beta][\alpha].$$

(4) $[\alpha][\beta] = [(\alpha * \mathbf{1})(\mathbf{1} * \beta)] = [\alpha * \beta] = [\alpha] * [\beta]$. □

As a consequence of (4), we can calculate the product of homotopy classes as a pointwise product of representatives and obtain:

Proposition C.1.9. (Hilton's Lemma) *For each topological group G, the fundamental group $\pi_1(G) := \pi_1(G, \mathbf{1})$ is abelian.*

Proof. We only have to combine (3) and (4) in Lemma C.1.8 for loops $\alpha, \beta \in \Omega(G, \mathbf{1})$. □

Exercises for Section C.1

Exercise C.1.1. If $f : X \to Y$ is a map between topological spaces and

$$X = X_1 \cup \ldots \cup X_n$$

holds with closed subsets $X_1, \ldots, X_n$, then f is continuous if and only if all restrictions $f|_{X_i}$ are continuous.

Exercise C.1.2. Show that the homotopy relation on $P(X, x_0, x_1)$ is an equivalence relation. Hint: Exercise C.1.1 helps to glue homotopies.

Exercise C.1.3. Show that, for $n > 1$, the sphere $\mathbb{S}^n$ is simply connected. For the proof, proceed along the following steps:

(a) Let $\gamma : [0, 1] \to \mathbb{S}^n$ be continuous. Then there exists an $m \in \mathbb{N}$ such that $\|\gamma(t) - \gamma(t')\| < \frac{1}{2}$ for $|t - t'| < \frac{1}{m}$.

(b) Define $\widetilde{\alpha} : [0, 1] \to \mathbb{R}^{n+1}$ as the piecewise affine curve with $\widetilde{\alpha}(\frac{k}{m}) = \gamma(\frac{k}{m})$ for $k = 0, \ldots, m$. Then $\alpha(t) := \frac{1}{\|\widetilde{\alpha}(t)\|}\widetilde{\alpha}(t)$ defines a continuous curve $\alpha : [0, 1] \to \mathbb{S}^n$.

(c) $\alpha \sim \gamma$. Hint: Consider $H(t,s) := \frac{(1-s)\gamma(t)+s\alpha(t)}{\|(1-s)\gamma(t)+s\alpha(t)\|}$.

(d) α is not surjective. The image of α is the central projection of a polygonal arc on the sphere.

(e) If $\beta \in \Omega(\mathbb{S}^n, y_0)$ is not surjective, then $\beta \sim y_0$ (it is homotopic to a constant map). Hint: Let $p \in \mathbb{S}^n \setminus \operatorname{im}\beta$. Using stereographic projection, where p corresponds to the point at infinity, show that $\mathbb{S}^n \setminus \{p\}$ is homeomorphic to $\mathbb{R}^n$, hence contractible.

(f) $\pi_1(\mathbb{S}^n, y_0) = \{[y_0]\}$ for $n \geq 2$ and $y_0 \in \mathbb{S}^n$.

Exercise C.1.4. Let X be a topological space, $x_0, x_1 \in X$ and $\alpha \in P(X, x_0, x_1)$ a path from x_0 to x_1. Show that the map

$$C\colon \pi_1(X, x_1) \to \pi_1(X, x_0), \quad [\gamma] \mapsto [\alpha * \gamma * \overline{\alpha}]$$

is an isomorphism of groups. In this sense the fundamental group does not depend on the base point if X is arcwise connected.

Exercise C.1.5. Let $\sigma\colon G \times X \to X$ be a continuous action of the topological group G on the topological space X and $x_0 \in X$. Then the orbit map $\sigma^{x_0}\colon G \to X, g \mapsto \sigma(g, x_0)$ defines a group homomorphism

$$\pi_1(\sigma^{x_0})\colon \pi_1(G) \to \pi_1(X, x_0).$$

Show that the image of this homomorphism is central, i.e., lies in the center of $\pi_1(X, x_0)$. Hint: Mimic the argument in the proof of Lemma C.1.8.

C.2 Coverings

In this section we discuss the concept of a covering map. One of its main applications is that it provides a means to calculate fundamental groups in terms of suitable coverings.

Definition C.2.1. Let X and Y be topological spaces. A continuous map $q\colon X \to Y$ is called a *covering* if each $y \in Y$ has an open neighborhood U such that $q^{-1}(U)$ is a non-empty disjoint union of open subsets $(V_i)_{i \in I}$, such that for each $i \in I$ the restriction $q|_{V_i}\colon V_i \to U$ is a homeomorphism. We call any such U an *elementary* open subset of X.

Note that this condition implies in particular that q is surjective and that the fibers of q are discrete subsets of X.

Examples C.2.2.

(a) The exponential function $\exp\colon \mathbb{C} \to \mathbb{C}^\times, z \mapsto e^z$ is a covering map.

(b) The map $q\colon \mathbb{R} \to \mathbb{T}, x \mapsto e^{ix}$ is a covering.

(c) The power maps $p_k\colon \mathbb{C}^\times \to \mathbb{C}^\times, z \mapsto z^k$ are coverings.

(d) If $q\colon G \to H$ is a surjective continuous open homomorphism of topological groups with discrete kernel, then q is a covering (Exercise C.2.2). All the examples (a)-(c) are of this type.

Lemma C.2.3. (Lebesgue number) [1] *Let (X, d) be a compact metric space and $(U_i)_{i\in I}$ an open cover. Then there exists a positive number $\lambda > 0$, called a Lebesgue number of the covering, such that any subset $S \subseteq X$ with diameter $\leq \lambda$ is contained in some U_i.*

Proof. Let us assume that such a number λ does not exist. Then there exists for each $n \in \mathbb{N}$ a subset S_n of diameter $\leq \frac{1}{n}$ which is not contained in some U_i. Pick a point $s_n \in S_n$. Then the sequence (s_n) has a subsequence converging to some $s \in X$. Then s is contained in some U_i, and since U_i is open, there exists an $\varepsilon > 0$ with $U_\varepsilon(s) \subseteq U_i$. If $n \in \mathbb{N}$ is such that $\frac{1}{n} < \frac{\varepsilon}{2}$ and $d(s_n, s) < \frac{\varepsilon}{2}$, we arrive at the contradiction $S_n \subseteq U_{\varepsilon/2}(s_n) \subseteq U_\varepsilon(s) \subseteq U_i$. □

Remark C.2.4. (1) If $(U_i)_{i\in I}$ is an open cover of the unit interval $[0, 1]$, then there exists an $n > 0$ such that all subsets of the form $\left[\frac{k}{n}, \frac{k+1}{n}\right]$, $k = 0, \ldots, n-1$, are contained in some U_i.

(2) If $(U_i)_{i\in I}$ is an open cover of the unit square $[0, 1]^2$, then there exists an $n > 0$ such that all subsets of the form

$$\left[\frac{k}{n}, \frac{k+1}{n}\right] \times \left[\frac{j}{n}, \frac{j+1}{n}\right], \quad k, j = 0, \ldots, n-1,$$

are contained in some U_i.

Theorem C.2.5. (The Path Lifting Property) *Let $q\colon X \to Y$ be a covering map and $\gamma\colon [0, 1] \to Y$ a path. Let $x_0 \in X$ be such that $q(x_0) = \gamma(0)$. Then there exists a unique path $\widetilde{\gamma}\colon [0, 1] \to X$ such that*

$$q \circ \widetilde{\gamma} = \gamma \quad \text{and} \quad \widetilde{\gamma}(0) = x_0.$$

Proof. Cover Y by elementary open set $U_i, i \in I$. By Remark C.2.4, applied to the open covering of I by the sets $\gamma^{-1}(U_i)$, there exists an $n \in \mathbb{N}$ such that all sets $\gamma(\left[\frac{k}{n}, \frac{k+1}{n}\right])$, $k = 0, \ldots, n-1$, are contained in some U_i. We now use induction to construct $\widetilde{\gamma}$. Let $V_0 \subseteq q^{-1}(U_0)$ be an open subset containing x_0 for which $q|_{V_0}$ is a homeomorphism onto U_0 and define $\widetilde{\gamma}$ on $\left[0, \frac{1}{n}\right]$ by

$$\widetilde{\gamma}(t) := (q|_{V_0})^{-1} \circ \gamma(t).$$

Assume that we have already constructed a continuous lift $\widetilde{\gamma}$ of γ on the interval $\left[0, \frac{k}{n}\right]$ and that $k < n$. Then we pick an elementary open subset U_i containing $\gamma([\frac{k}{n}, \frac{k+1}{n}])$ and an open subset $V_k \subseteq X$ containing $\widetilde{\gamma}(\frac{k}{n})$ for which $q|_{V_k}$ is a homeomorphism onto U_i. We then define $\widetilde{\gamma}$ for $t \in \left[\frac{k}{n}, \frac{k+1}{n}\right]$ by

$$\widetilde{\gamma}(t) := (q|_{V_k})^{-1} \circ \gamma(t).$$

We thus obtain the required lift $\widetilde{\gamma}$ of γ on $[0, \frac{k}{n+1}]$.

If $\widehat{\gamma}\colon [0, 1] \to X$ is any continuous lift of γ with $\widehat{\gamma}(0) = x_0$, then $\widehat{\gamma}(\left[0, \frac{1}{n}\right])$ is a connected subset of $q^{-1}(U_0)$ containing x_0, hence contained in V_0, showing that $\widetilde{\gamma}$ coincides with $\widehat{\gamma}$ on $\left[0, \frac{1}{n}\right]$. Applying the same argument at each step of the induction, we obtain $\widehat{\gamma} = \widetilde{\gamma}$, so that the lift $\widetilde{\gamma}$ is unique. □

[1] Lebesgue, Henri (1875–1941)

Theorem C.2.6. (The Covering Homotopy Theorem) *Let* $I := [0,1]$ *and* $q\colon X \to Y$ *be a covering map and* $H\colon I^2 \to Y$ *be a homotopy with fixed endpoints of the paths* $\gamma := H_0$ *and* $\eta := H_1$*. For any lift* $\widetilde{\gamma}$ *of* γ *there exists a unique lift* $G\colon I^2 \to X$ *of* H *with* $G_0 = \widetilde{\gamma}$*. Then* $\widetilde{\eta} := G_1$ *is the unique lift of* η *starting in the same point as* $\widetilde{\gamma}$ *and* G *is a homotopy from* $\widetilde{\gamma}$ *to* $\widetilde{\eta}$*. In particular, lifts of homotopic curves in* Y *starting in the same point are homotopic in* X*.*

Proof. Using the Path Lifting Property (Theorem C.2.5), we find for each $t \in I$ a unique continuous lift $I \to X, s \mapsto G(s,t)$, starting in $\widetilde{\gamma}(t)$ with $q(G(s,t)) = H(s,t)$. It remains to show that the map $G\colon I^2 \to X$ obtained in this way is continuous.

So let $s \in I$. Using Remark C.2.4, we find a natural number n such that for each connected neighborhood W_s of s of diameter $\leq \frac{1}{n}$ and each $i = 0,\ldots,n$, the set $H\big(W_s \times \big[\frac{k}{n}, \frac{k+1}{n}\big]\big)$ is contained in some elementary subset U_k of Y. Assuming that G is continuous in $W_s \times \{\frac{k}{n}\}$, G maps this set into a connected subset of $q^{-1}(U_k)$, hence into some open subset V_k for which $q|_{V_k}$ is a homeomorphism onto U_k. But then the lift G on $W_s \times \big[\frac{k}{n}, \frac{k+1}{n}\big]$ must be contained in V_k, so that it is of the form $(q|_{V_k})^{-1} \circ H$, hence continuous. This means that G is continuous on $W_s \times \big[\frac{k}{n}, \frac{k+1}{n}\big]$. Now an inductive argument shows that G is continuous on $W_s \times I$ and hence on the whole square I^2.

Since the fibers of q are discrete and the curves $s \mapsto H(s,0)$ and $s \mapsto H(s,1)$ are constant, the curves $G(s,0)$ and $G(s,1)$ are also constant. Therefore $\widetilde{\eta}$ is the unique lift of η starting in $\widetilde{\gamma}(0) = G(0,0) = G(1,0)$ and G is a homotopy with fixed endpoints from $\widetilde{\gamma}$ to $\widetilde{\eta}$. □

Corollary C.2.7. *If* $q\colon X \to Y$ *is a covering with* $q(x_0) = y_0$*, then the corresponding group homomorphism*

$$\pi_1(q,x_0)\colon \pi_1(X,x_0) \to \pi_1(Y,y_0), \quad [\gamma] \mapsto [q \circ \gamma]$$

is injective.

Proof. If γ, η are loops in x_0 with $[q\circ\gamma] = [q\circ\eta]$, then the Covering Homotopy Theorem C.2.6 implies that γ and η are homotopic. Therefore $[\gamma] = [\eta]$ shows that $\pi_1(q,x_0)$ is injective. □

Corollary C.2.8. *If* Y *is simply connected and* X *is arcwise connected, then each covering map* $q\colon X \to Y$ *is a homeomorphism.*

Proof. Since q is an open continuous map, it remains to show that q is injective. So pick $x_0 \in X$ and $y_0 \in Y$ with $q(x_0) = y_0$. If $x \in X$ also satisfies $q(x) = y_0$, then there exists a path $\alpha \in P(X,x_0,x)$ from x_0 to x. Now $q \circ \alpha$ is a loop in Y, hence contractible because Y is simply connected. Now the Covering Homotopy Theorem implies that the unique lift α of $q \circ \alpha$ starting in x_0 is a loop, and therefore that $x_0 = x$. This proves that q is injective. □

The following theorem provides a more powerful tool, from which the preceding corollary easily follows. We recall that a topological space X is called *locally arcwise connected* if each neighborhood U of a point $x \in X$ contains some arcwise connected neighborhood V of x (cf. Exercise A.5.5).

Theorem C.2.9. (The Lifting Theorem) *Assume that $q\colon X \to Y$ is a covering map with $q(x_0) = y_0$, that W is arcwise connected and locally arcwise connected, and that $f\colon W \to Y$ is a given map with $f(w_0) = y_0$. Then a continuous map $g\colon W \to X$ with*

$$g(w_0) = x_0 \quad \textit{and} \quad q \circ g = f \tag{C.1}$$

exists if and only if

$$\pi_1(f, w_0)(\pi_1(W, w_0)) \subseteq \pi_1(q, x_0)(\pi_1(X, x_0)), \ \textit{i.e.}\ \operatorname{im}(\pi_1(f, w_0)) \subseteq \operatorname{im}(\pi_1(q, x_0)). \tag{C.2}$$

If g exists, which is always the case if W is simply connected, then it is uniquely determined by (C.1)*. Condition* (C.2) *is in particular satisfied if W is simply connected.*

Proof. If g exists, then $f = q \circ g$ implies that the image of the homomorphism $\pi_1(f, w_0) = \pi_1(q, x_0) \circ \pi_1(g, w_0)$ is contained in the image of $\pi_1(q, x_0)$.

Let us, conversely, assume that this condition is satisfied. To define g, let $w \in W$ and $\alpha_w\colon I \to W$ be a path from w_0 to w. Then $f \circ \alpha_w\colon I \to Y$ is a path which has a continuous lift $\beta_w\colon I \to X$ starting in x_0. We claim that $\beta_w(1)$ does not depend on the choice of the path α_w. Indeed, if α'_w is another path from w_0 to w, then $\alpha_w * \overline{\alpha'_w}$ is a loop in w_0, so that $(f \circ \alpha_w) * (f \circ \overline{\alpha'_w})$ is a loop in y_0. In view of (C.2), the homotopy class of this loop is contained in the image of $\pi_1(q, x_0)$, so that it has a lift $\eta\colon I \to X$ which is a loop in x_0. Since the reverse of the second half $\eta|_{[\frac{1}{2},1]}$ of η is a lift of $f \circ \alpha'_w$, starting in x_0, it is β'_w, or, more precisely

$$\beta'_w(t) = \eta\Big(1 - \frac{t}{2}\Big) \quad \text{for} \quad 0 \le t \le 1.$$

We thus obtain

$$\beta'_w(1) = \eta\Big(\frac{1}{2}\Big) = \beta_w(1).$$

We now put $g(w) := \beta_w(1)$, and it remains to see that g is continuous. This is where we shall use the assumption that W is locally arcwise connected. Let $w \in W$ and put $y := f(w)$. Further, let $U \subseteq Y$ be an elementary neighborhood of y and V be an arcwise connected neighborhood of w in W such that $f(V) \subseteq U$. Fix a path α_w from w_0 to w as before. For any point $w' \in W$ we choose a path $\gamma_{w'}$ from w to w' in V, so that $\alpha_w * \gamma_{w'}$ is a path from w_0 to w'. Let $\widetilde{U} \subseteq X$ be an open subset of X for which $q|_{\widetilde{U}}$ is a homeomorphism onto U and $g(w) \in \widetilde{U}$. Then the uniqueness of lifts implies that

$$\beta_{w'} = \beta_w * \big((q|_{\widetilde{U}})^{-1} \circ (f \circ \gamma_{w'})\big).$$

We conclude that

$$g(w') = (q|_{\widetilde{U}})^{-1}(f(w')) \in \widetilde{U},$$

hence that $g|_V$ is continuous.

We finally show that g is unique. In fact, if $h\colon W \to X$ is another lift of f satisfying $h(w_0) = x_0$, then the set $S := \{w \in W\colon g(w) = h(w)\}$ is non-empty and closed. We claim that it is also open. In fact, let $w_1 \in S$ and U be a connected open elementary neighborhood of $f(w_1)$ and V an arcwise connected neighborhood of w_1 with $f(V) \subseteq U$. If $\widetilde{U} \subseteq q^{-1}(U)$ is the open subset on which q is a homeomorphism containing

$g(w_1) = h(w_1)$, then the arcwise connectedness of V implies that $g(V), h(V) \subseteq \widetilde{U}$, and hence that $V \subseteq S$. Therefore S is open, closed and non-empty, so that the connectedness of W yields $S = W$, i.e., $g = h$. □

Corollary C.2.10. (Uniqueness of Simply Connected Coverings) *Suppose that Y is locally arcwise connected. If $q_1 \colon X_1 \to Y$ and $q_2 \colon X_2 \to Y$ are two simply connected arcwise connected coverings, then there exists a homeomorphism $\varphi \colon X_1 \to X_2$ with $q_2 \circ \varphi = q_1$.*

Proof. Since Y is locally arcwise connected, both covering spaces X_1 and X_2 also have this property. Pick points $x_1 \in X_1$, $x_2 \in X_2$ with $y := q_1(x_1) = q_2(x_2)$. According to the Lifting Theorem C.2.9, there exists a unique lift $\varphi \colon X_1 \to X_2$ of q_1 with $\varphi(x_1) = x_2$. We likewise obtain a unique lift $\psi \colon X_2 \to X_1$ of q_2 with $\psi(x_2) = x_1$. Then $\varphi \circ \psi \colon X_1 \to X_1$ is a lift of id_Y fixing x_1, so that the uniqueness of lifts implies that $\varphi \circ \psi = \mathrm{id}_{X_1}$. The same argument yields $\psi \circ \varphi = \mathrm{id}_{X_2}$, so that φ is a homeomorphism with the required properties. □

Definition C.2.11. A topological space X is called *semilocally simply connected* if each point $x_0 \in X$ has a neighborhood U such that each loop $\alpha \in \Omega(U, x_0)$ is homotopic to $[x_0]$ in X, i.e., the natural homomorphism

$$\pi(i_U) \colon \pi_1(U, x_0) \to \pi_1(X, x_0), \quad [\gamma] \mapsto [i_U \circ \gamma]$$

induced by the inclusion map $i_U \colon U \to X$ is trivial.

Example C.2.12. Every manifold M is locally arcwise connected and semilocally simply connected. In fact, every neighborhood U of a point $m \in M$ contains an open neighborhood V homeomorphic to an open ball B in a Banach space E. Since B is convex, it is arcwise connected and simply connected.

Theorem C.2.13. *Let Y be arcwise connected and locally arcwise connected. Then Y has a simply connected covering space if and only if Y is semilocally simply connected.*

Proof. If $q \colon X \to Y$ is a simply connected covering space and $U \subseteq Y$ is a pathwise connected elementary open subset. Then each loop γ in U lifts to a loop $\widetilde{\gamma}$ in X, and since $\widetilde{\gamma}$ is homotopic to a constant map in X, the same holds for the loop $\gamma = q \circ \widetilde{\gamma}$ in Y.

Conversely, let us assume that Y is semilocally simply connected. We choose a base point $y_0 \in Y$ and let

$$\widetilde{Y} := P(Y, y_0)/\sim$$

be the set of homotopy classes of paths starting in y_0. We shall topologize $\widetilde{Y}$ in such a way that the map

$$q \colon \widetilde{Y} \to Y, \quad [\gamma] \mapsto \gamma(1)$$

defines a simply connected covering of Y.

Let $\mathcal{B}$ denote the set of all arcwise connected open subsets $U \subseteq Y$ for which each loop in U is contractible in Y and note that our assumptions on Y imply that $\mathcal{B}$ is a basis for the topology of Y, i.e., each open subset is a union of elements of $\mathcal{B}$. If $\gamma \in P(Y, y_0)$ satisfies $\gamma(1) \in U \in \mathcal{B}$, let

$$U_{[\gamma]} := \{[\eta] \in q^{-1}(U) \colon (\exists \beta \in C(I, U))\ \eta \sim \gamma * \beta\}.$$

We shall now verify several properties of these definitions, culminating in the proof of the theorem.

(1) $[\eta] \in U_{[\gamma]} \Rightarrow U_{[\eta]} = U_{[\gamma]}$.

To prove this, let $[\zeta] \in U_{[\eta]}$. Then $\zeta \sim \eta * \beta$ for some path β in U. Further $\eta \sim \gamma * \beta'$ for some path β' in U. Now $\zeta \sim \gamma * \beta' * \beta$, and $\beta' * \beta$ is a path in U, so that $[\zeta] \in U_{[\gamma]}$. This proves $U_{[\eta]} \subseteq U_{[\gamma]}$. We also have $\gamma \sim \eta * \overline{\beta'}$, so that $[\gamma] \in U_{[\eta]}$, and the first part implies that $U_{[\gamma]} \subseteq U_{[\eta]}$.

(2) q maps $U_{[\gamma]}$ injectively onto U.

That $q(U_{[\gamma]}) = U$ is clear since U and Y are arcwise connected. To show that that it is one-to-one, let $[\eta], [\eta'] \in U_{[\gamma]}$, which we know from (1) is the same as $U_{[\eta]}$. Suppose $\eta(1) = \eta'(1)$. Since $[\eta'] \in U_{[\eta]}$, we have $\eta' \sim \eta * \alpha$ for some loop α in U. But then α is contractible in Y, so that $\eta' \sim \eta$, i.e., $[\eta'] = [\eta]$.

(3) $U, V \in \mathcal{B}$, $\gamma(1) \in U \subseteq V$, implies $U_{[\gamma]} \subseteq V_{[\gamma]}$.

This is trivial.

(4) The sets $U_{[\gamma]}$ for $U \in \mathcal{B}$ and $[\gamma] \in \widetilde{Y}$ form a basis of a topology on $\widetilde{Y}$.

Suppose $[\gamma] \in U_{[\eta]} \cap V_{[\eta']}$. Let $W \subseteq U \cap V$ be in $\mathcal{B}$ with $\gamma(1) \in W$. Then $[\gamma] \in W_{[\gamma]} \subseteq U_{[\gamma]} \cap V_{[\gamma]} = U_{[\eta]} \cap V_{[\eta']}$.

(5) q is open and continuous.

We have already seen in (2) that $q(U_{[\gamma]}) = U$, and these sets form a basis of the topology on $\widetilde{Y}$, resp., Y. Therefore q is an open map. We also have for $U \in \mathcal{B}$ the relation

$$q^{-1}(U) = \bigcup_{\gamma(1) \in U} U_{[\gamma]},$$

which is open. Hence q is continuous.

(6) $q|_{U_{[\gamma]}}$ is a homeomorphism.

This is because it is bijective, continuous and open.

At this point we have shown that $q \colon \widetilde{Y} \to Y$ is a covering map. It remains to see that $\widetilde{Y}$ is arcwise connected and simply connected.

(7) Let $H \colon I \times I \to Y$ be a continuous map with $H(t, 0) = y_0$. Then $h_t(s) := H(t, s)$ defines a path in Y starting in y_0. Let $\widetilde{h}(t) := [h_t] \in \widetilde{Y}$. Then $\widetilde{h}$ is a path in $\widetilde{Y}$ covering the path $t \mapsto h_t(1) = H(t, 1)$ in Y. We claim that $\widetilde{h}$ is continuous. Let $t_0 \in I$. We shall prove continuity at t_0. Let $U \in \mathcal{B}$ be a neighborhood of $h_{t_0}(1)$. Then there exists an interval $I_0 \subseteq I$ which is a neighborhood of t_0 with $h_t(1) \in U$ for $t \in I_0$. Then $\alpha(s) := H(t_0 + s(t - t_0), 1)$ is a continuous curve in U with $\alpha(0) = h_{t_0}(1)$ and $\alpha(1) = h_t(1)$, so that $h_{t_0} * \alpha$ is curve with the same endpoint as h_t. Applying Exercise C.2.1 to the restriction of H to the interval between t_0 and t, we see that $h_t \sim h_{t_0} * \alpha$, so that $\widetilde{h}(t) = [h_t] \in U_{[h_{t_0}]}$ for $t \in I_0$. Since $q|_{U_{[h_{t_0}]}}$ is a homeomorphism, $\widetilde{h}$ is continuous in t_0.

(8) $\widetilde{Y}$ is arcwise connected.

For $[\gamma] \in \widetilde{Y}$ put $h_t(s) := \gamma(st)$. By (7), this yields a path $\widetilde{\gamma}(t) = [h_t]$ in $\widetilde{Y}$ from $\widetilde{y}_0 := [y_0]$ (the class of the constant path) to the point $[\gamma]$.

(9) $\widetilde{Y}$ is simply connected.

Let $\widetilde{\alpha} \in \Omega(\widetilde{Y}, \widetilde{y}_0)$ be a loop in $\widetilde{Y}$ and $\alpha := q \circ \widetilde{\alpha}$ its image in Y. Let $h_t(s) := \alpha(st)$. Then we have the path $\widetilde{h}(t) = [h_t]$ in $\widetilde{Y}$ from (7). This path covers α since $h_t(1) = \alpha(t)$. Further, $\widetilde{h}(0) = \widetilde{y}_0$ is the constant path. Also, by definition, $\widetilde{h}(1) = [\alpha]$. From the uniqueness of lifts we derive that $\widetilde{h} = \widetilde{\alpha}$ is closed, so that $[\alpha] = [y_0]$. Therefore the homomorphism

$$\pi_1(q, y_0) \colon \pi_1(\widetilde{Y}, \widetilde{y}_0) \to \pi_1(Y, y_0)$$

vanishes. Since it is also injective (Corollary C.2.7), $\pi_1(\widetilde{Y}, \widetilde{y}_0)$ is trivial, i.e., $\widetilde{Y}$ is simply connected. □

Definition C.2.14. Let $q \colon X \to Y$ be a covering. A homeomorphism $\varphi \colon X \to X$ is called a *deck transformation* of the covering if $q \circ \varphi = q$. This means that φ permutes the elements in the fibers of q. We write $\mathrm{Deck}(X, q)$ for the group of deck transformations.

Example C.2.15. For the covering map $\exp \colon \mathbb{C} \to \mathbb{C}^\times$, the deck transformations have the form

$$\varphi(z) = z + 2\pi i n, \quad n \in \mathbb{Z}.$$

Remark C.2.16. Suppose that X is connected and $\varphi, \psi \colon X \to X$ are deck transformations with $\varphi(x_0) = \psi(x_0)$ for some $x_0 \in X$. Then $q \circ \varphi = q = q \circ \psi$ and the uniqueness assertion of the Lifting Theorem C.2.9 imply that $\varphi = \psi$. This means that deck transformations are determined by the image of x_0, resp., the map

$$\mathrm{ev}_{x_0} \colon \mathrm{Deck}(X, q) \to X, \quad \varphi \mapsto \varphi(x_0)$$

is injective. If, in addition, X is simply connected, the Lifting Theorem C.2.9 further implies that $\mathrm{im}(\mathrm{ev}_{x_0}) = q^{-1}(q(x_0))$.

Proposition C.2.17. *Let $q \colon \widetilde{Y} \to Y$ be a simply connected covering of the connected locally arcwise connected space Y. Pick $\widetilde{y}_0 \in \widetilde{Y}$ and put $y_0 := q(\widetilde{y}_0)$. For each $[\gamma] \in \pi_1(Y, y_0)$ we write $\varphi_{[\gamma]} \in \mathrm{Deck}(\widetilde{Y}, q)$ for the unique lift of id_X mapping $\widetilde{y}_0$ to the endpoint $\widetilde{\gamma}(1)$ of the lift $\widetilde{\gamma}$ of γ starting in $\widetilde{y}_0$. Then the map*

$$\Phi \colon \pi_1(Y, y_0) \to \mathrm{Deck}(\widetilde{Y}, q), \quad \Phi([\gamma]) = \varphi_{[\gamma]}$$

is an isomorphism of groups.

Proof. For $\gamma, \eta \in \Omega(Y, y_0)$, the composition $\varphi_{[\gamma]} \circ \varphi_{[\eta]}$ is a deck transformation mapping $\widetilde{y}_0$ to the endpoint of $\varphi_{[\gamma]} \circ \widetilde{\eta}$ which coincides with the endpoint of the lift of η starting in $\widetilde{\gamma}(1)$. Hence it also is the endpoint of the lift of the loop $\gamma * \eta$. This leads to $\varphi_{[\gamma]} \circ \varphi_{[\eta]} = \varphi_{[\gamma * \eta]}$, so that Φ is a group homomorphism.

To see that Φ is injective, we note that $\varphi_{[\gamma]} = \mathrm{id}_{\widetilde{Y}}$ implies that $\widetilde{\gamma}(1) = \widetilde{y}_0$, so that $\widetilde{\gamma}$ is a loop, and hence that $[\gamma] = [y_0]$.

For the surjectivity, let φ be a deck transformation and $y := \varphi(\widetilde{y}_0)$. If α is a path from $\widetilde{y}_0$ to y, then $\gamma := q \circ \alpha$ is a loop in y_0 with $\alpha = \widetilde{\gamma}$, so that $\varphi_{[\gamma]}(\widetilde{y}_0) = y$, and the uniqueness of lifts (Theorem C.2.9) implies that $\varphi = \varphi_{[\gamma]}$. $\square$

Example C.2.18. With Example C.2.15 and the simple connectedness of $\mathbb{C}$ we derive that

$$\pi_1(\mathbb{C}^\times, 1) \cong \mathrm{Deck}(\mathbb{C}, \exp) \cong \mathbb{Z}.$$

Exercises for Section C.2

Exercise C.2.1. Let $F\colon I^2 \to X$ be a continuous map with $F(0,s) = x_0$ for $s \in I$ and define

$$\gamma(t) := F(t,0), \quad \eta(t) := F(t,1), \quad \alpha(t) := F(1,t), \quad t \in I.$$

Show that $\gamma * \alpha \sim \eta$. Hint: Consider the map

$$G\colon I^2 \to I^2, \quad G(t,s) := \begin{cases} (2t, s) & \text{for } 0 \leq t \leq \frac{1}{2}, s \leq 1 - 2t, \\ (1, 2t-1) & \text{for } \frac{1}{2} \leq t \leq 1, s \leq 2t-1, \\ (t + \frac{1-s}{2}, s) & \text{else} \end{cases}$$

and show that it is continuous. Take a look at the boundary values of $F \circ G$.

Exercise C.2.2. Let $q\colon G \to H$ be an morphism of topological groups with discrete kernel Γ. Show that:

(1) If $V \subseteq G$ is an open $\mathbf{1}$-neighborhood with $(V^{-1}V) \cap \Gamma = \{\mathbf{1}\}$ and q is open, then $q|_V : V \to q(V)$ is a homeomorphism.

(2) If q is open and surjective, then q is a covering.

(3) If q is open and H is connected, then q is surjective, hence a covering.

Exercise C.2.3. A map $f\colon X \to Y$ between topological spaces is called a *local homeomorphism* if each point $x \in X$ has an open neighborhood U such that $f|_U : U \to f(U)$ is a homeomorphism onto an open subset of Y.

(1) Show that each covering map is a local homeomorphism.

(2) Find a surjective local homeomorphism which is not a covering. Can you also find an example where X is connected?

Exercise C.2.4. In the euclidean plane $\mathbb{R}^2$, we write

$$C_r(m) := \{x \in \mathbb{R}^2 \colon \|x - m\|_2 = r\}$$

for the circle of radius r and center m. Consider the union

$$X := \bigcup_{n \in \mathbb{N}} C_{1/n}\Big(\frac{1}{n}, 0\Big).$$

Show that X is arcwise connected but not semilocally simply connected. Hint: Consider the point $(0,0) \in X$.

Exercise C.2.5. Let $\mathbb{T} \subseteq \mathbb{C}$ be the unit circle. Show that:

(i) For every continuous map $f\colon \mathbb{T} \to \mathbb{T}$, there exists a continuous map $\beta\colon \mathbb{R} \to \mathbb{R}$ with $f(e^{it}) = e^{i\beta(t)}$ for $t \in \mathbb{R}$.

(ii) Show that for every map β as in (i),

$$\deg(f) := \frac{\beta(t + 2\pi) - \beta(t)}{2\pi} \in \mathbb{Z}$$

is constant. It is called the *winding number* or *mapping degree* of f. Verify that this number does not depend on the choice of β for a given function f.

(iii) Two maps $f_1, f_2\colon \mathbb{T} \to \mathbb{T}$ are homotopic if and only if $\deg(f_1) = \deg(f_2)$. Hint: Use the Covering Homotopy Theorem to see that this condition is necessary. To see that it is sufficient, proceed with the ansatz $\beta_s := (1 - s)\beta_0 + s\beta_1$ and show that this defines a homotopy by $f_s(e^{it}) = e^{i\beta_s(t)}$.

(iv) Show that the same arguments are equally valid for maps $f\colon \mathbb{C}^\times \to \mathbb{C}^\times$.

Bibliography

[HiNe91] Hilgert, J., and K.-H. Neeb, "Lie-Gruppen und Lie-Algebren", Vieweg, 1991

[La93] Lang, S., "Algebra," Addison Wesley Publ. Comp., London, 1993

[La99] —, "Fundamentals of Differential Geometry", Grad. Texts Math. **191**, Springer-Verlag, Berlin, 1999

[Ne04] Neeb, K.-H., *Infinite-dimensional Lie groups and their representations*, in "Lie Theory: Lie Algebras and Representations," Progress in Math. **228**, Ed. J. P. Anker, B. Ørsted, Birkhäuser Verlag, 2004; 213–328

[Ru73] Rudin, W., "Functional Analysis," McGraw Hill, 1973

[Ru86] Rudin, W., "Real and Complex Analysis," McGraw Hill, 1986

[St94] Sternberg, S., "Group Theory and Physics," Cambridge University Press, 1994 (A nice book discussing many application to concrete problems from physics and chemistry.)