*Chapter 1*

# Introduction

This text focuses on three related topics:

- *Hilbert's fifth problem* on the topological description of Lie groups, as well as the closely related (local) classification of *locally compact groups* (the *Gleason-Yamabe theorem*, see Theorem 1.1.13);
- Approximate groups in nonabelian groups, and their classification [**Hr2012**], [**BrGrTa2011**] via the Gleason-Yamabe theorem; and
- Gromov's theorem [**Gr1981**] on groups of polynomial growth, as proven via the classification of approximate groups (as well as some consequences to fundamental groups of Riemannian manifolds).

These three families of results exemplify two broad principles (part of what I like to call the *the dichotomy between structure and randomness* [**Ta2008**]):

- (Rigidity) If a group-like object exhibits a weak amount of regularity, then it (or a large portion thereof) often automatically exhibits a strong amount of regularity as well.
- (Structure) Furthermore, this strong regularity manifests itself either as Lie type structure (in continuous settings) or *nilpotent* type structure (in discrete settings). (In some cases, "nilpotent" should be replaced by sister properties such as "*abelian*", "*solvable*", or "*polycyclic*".)

Let us illustrate these two principles with two simple examples, one in the continuous setting and one in the discrete setting. We begin with a continuous example. Given an $n \times n$ complex matrix $A \in M_n(\mathbf{C})$, define the

*matrix exponential* $\exp(A)$ of $A$ by the formula

$$\exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!} = 1 + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots$$

which can easily be verified to be an absolutely convergent series.

**Exercise 1.0.1.** Show that the map $A \mapsto \exp(A)$ is a real analytic (and even complex analytic) map from $M_n(\mathbf{C})$ to $M_n(\mathbf{C})$, and obeys the restricted homomorphism property

$$\exp(sA)\exp(tA) = \exp((s+t)A) \tag{1.1}$$

for all $A \in M_n(\mathbf{C})$ and $s, t \in \mathbf{C}$.

**Proposition 1.0.1** (Rigidity and structure of matrix homomorphisms)**.** *Let $n$ be a natural number. Let $\mathrm{GL}_n(\mathbf{C})$ be the group of invertible $n \times n$ complex matrices. Let $\Phi : \mathbf{R} \to \mathrm{GL}_n(\mathbf{C})$ be a map obeying two properties:*

(1) *(Group-like object) $\Phi$ is a homomorphism, thus $\Phi(s)\Phi(t) = \Phi(s+t)$ for all $s, t \in \mathbf{R}$.*

(2) *(Weak regularity) The map $t \mapsto \Phi(t)$ is continuous.*

*Then:*

(i) *(Strong regularity) The map $t \mapsto \Phi(t)$ is smooth (i.e., infinitely differentiable). In fact it is even real analytic.*

(ii) *(Lie-type structure) There exists a (unique) complex $n \times n$ matrix $A$ such that $\Phi(t) = \exp(tA)$ for all $t \in \mathbf{R}$.*

**Proof.** Let $\Phi$ be as above. Let $\varepsilon > 0$ be a small number (depending only on $n$). By the homomorphism property, $\Phi(0) = 1$ (where we use 1 here to denote the identity element of $\mathrm{GL}_n(\mathbf{C})$), and so by continuity we may find a small $t_0 > 0$ such that $\Phi(t) = 1 + O(\varepsilon)$ for all $t \in [-t_0, t_0]$ (we use some arbitrary norm here on the space of $n \times n$ matrices, and allow implied constants in the $O()$ notation to depend on $n$).

The map $A \mapsto \exp(A)$ is real analytic and (by the *inverse function theorem*) is a diffeomorphism near 0. Thus, by the inverse function theorem, we can (if $\varepsilon$ is small enough) find a matrix $B$ of size $B = O(\varepsilon)$ such that $\Phi(t_0) = \exp(B)$. By the homomorphism property and (1.1), we thus have

$$\Phi(t_0/2)^2 = \Phi(t_0) = \exp(B) = \exp(B/2)^2.$$

On the other hand, by another application of the inverse function theorem we see that the squaring map $A \mapsto A^2$ is a diffeomorphism near 1 in $\mathrm{GL}_n(\mathbf{C})$, and thus (if $\varepsilon$ is small enough)

$$\Phi(t_0/2) = \exp(B/2).$$

We may iterate this argument (for a fixed, but small, value of $\varepsilon$) and conclude that

$$\Phi(t_0/2^k) = \exp(B/2^k)$$

for all $k = 0, 1, 2, \ldots$. By the homomorphism property and (1.1) we thus have

$$\Phi(qt_0) = \exp(qB)$$

whenever $q$ is a dyadic rational, i.e., a rational of the form $a/2^k$ for some integer $a$ and natural number $k$. By continuity we thus have

$$\Phi(st_0) = \exp(sB)$$

for all real $s$. Setting $A := B/t_0$ we conclude that

$$\Phi(t) = \exp(tA)$$

for all real $t$, which gives existence of the representation and also real analyticity and smoothness. Finally, uniqueness of the representation $\Phi(t) = \exp(tA)$ follows from the identity

$$A = \frac{d}{dt} \exp(tA)|_{t=0}. \qquad \square$$

**Exercise 1.0.2.** Generalise Proposition 1.0.1 by replacing the hypothesis that $\Phi$ is continuous with the hypothesis that $\Phi$ is Lebesgue measurable. (*Hint:* Use the *Steinhaus theorem*, see e.g. [**Ta2011**, Exercise 1.6.8].) Show that the proposition fails (assuming the axiom of choice) if this hypothesis is omitted entirely.

Note how one needs both the group-like structure and the weak regularity in combination in order to ensure the strong regularity; neither is sufficient on its own. We will see variants of the above basic argument throughout the course. Here, the task of obtaining smooth (or real analytic structure) was relatively easy, because we could borrow the smooth (or real analytic) structure of the domain $\mathbf{R}$ and range $M_n(\mathbf{C})$; but, somewhat remarkably, we shall see that one can still build such smooth or analytic structures even when none of the original objects have any such structure to begin with.

Now we turn to a second illustration of the above principles, namely *Jordan's theorem* [**Jo1878**], which uses a discreteness hypothesis to upgrade Lie type structure to nilpotent (and in this case, abelian) structure. We shall formulate Jordan's theorem in a slightly stilted fashion in order to emphasise the adherence to the above-mentioned principles.

**Theorem 1.0.2** (Jordan's theorem)**.** *Let $G$ be an object with the following properties:*

(1) *(Group-like object) $G$ is a group.*

(2) *(Discreteness) $G$ is finite.*

(3) *(Lie-type structure) $G$ is a subgroup of $\mathrm{U}_n(\mathbf{C})$ (the group of unitary $n \times n$ matrices) for some $n$.*

*Then there is a subgroup $G'$ of $G$ such that*

(i) *($G'$ is close to $G$) The index $|G/G'|$ of $G'$ in $G$ is $O_n(1)$ (i.e., bounded by $C_n$ for some quantity $C_n$ depending only on $n$).*

(ii) *(Nilpotent-type structure) $G'$ is abelian.*

A key observation in the proof of Jordan's theorem is that if two unitary elements $g, h \in \mathrm{U}_n(\mathbf{C})$ are close to the identity, then their *commutator* $[g,h] = g^{-1}h^{-1}gh$ is even closer to the identity (in, say, the operator norm $\|\|_{\mathrm{op}}$). Indeed, since multiplication on the left or right by unitary elements does not affect the operator norm, we have

$$\begin{aligned}\|[g,h]-1\|_{\mathrm{op}} &= \|gh-hg\|_{\mathrm{op}} \\ &= \|(g-1)(h-1)-(h-1)(g-1)\|_{\mathrm{op}}\end{aligned}$$

and so by the triangle inequality

$$\|[g,h]-1\|_{\mathrm{op}} \leq 2\|g-1\|_{\mathrm{op}}\|h-1\|_{\mathrm{op}}. \tag{1.2}$$

Now we can prove Jordan's theorem.

**Proof.** We induct on $n$, the case $n = 1$ being trivial. Suppose first that $G$ contains a *central element* $g$ (i.e., an element that commutes with every element in $G$) which is not a multiple of the identity. Then, by definition, $G$ is contained in the *centraliser* $Z(g) := \{h \in \mathrm{U}_n(\mathbf{C}) : gh = hg\}$ of $g$, which by the spectral theorem is isomorphic to a product $\mathrm{U}_{n_1}(\mathbf{C}) \times \cdots \times \mathrm{U}_{n_k}(\mathbf{C})$ of smaller unitary groups. Projecting $G$ to each of these factor groups and applying the induction hypothesis, we obtain the claim.

Thus we may assume that $G$ contains no central elements other than multiples of the identity. Now pick a small $\varepsilon > 0$ (one could take $\varepsilon = \frac{1}{10n}$ in fact) and consider the subgroup $G'$ of $G$ generated by those elements of $G$ that are within $\varepsilon$ of the identity (in the operator norm). By considering a maximal $\varepsilon$-net of $G$ we see that $G'$ has index at most $O_{n,\varepsilon}(1)$ in $G$. By arguing as before, we may assume that $G'$ has no central elements other than multiples of the identity.

If $G'$ consists only of multiples of the identity, then we are done. If not, take an element $g$ of $G'$ that is not a multiple of the identity, and which is as close as possible to the identity (here is where we crucially use that $G$ is finite). Note that $g$ is within $\varepsilon$ of the identity. By (1.2), we see that if $\varepsilon$ is sufficiently small depending on $n$, and if $h$ is one of the generators of $G'$, then $[g,h]$ lies in $G'$ and is closer to the identity than $g$, and is thus a

multiple of the identity. On the other hand, $[g,h]$ has determinant 1. Given that it is so close to the identity, it must therefore be the identity (if $\varepsilon$ is small enough). In other words, $g$ is central in $G'$, and is thus a multiple of the identity. But this contradicts the hypothesis that there are no central elements other than multiples of the identity, and we are done. □

Commutator estimates such as (1.2) will play a fundamental role in many of the arguments we will see in this text; as we saw above, such estimates combine very well with a discreteness hypothesis, but will also be very useful in the continuous setting.

**Exercise 1.0.3.** Generalise Jordan's theorem to the case when $G$ is a finite subgroup of $\mathrm{GL}_n(\mathbf{C})$ rather than of $\mathrm{U}_n(\mathbf{C})$. (*Hint:* The elements of $G$ are not necessarily unitary, and thus do not necessarily preserve the standard Hilbert inner product of $\mathbf{C}^n$. However, if one averages that inner product by the finite group $G$, one obtains a new inner product on $\mathbf{C}^n$ that is preserved by $G$, which allows one to conjugate $G$ to a subgroup of $\mathrm{U}_n(\mathbf{C})$. This averaging trick is (a small) part of *Weyl's unitary trick* in representation theory.)

**Remark 1.0.3.** We remark that one can strengthen Jordan's theorem further by relaxing the finiteness assumption on $G$ to a periodicity assumption; see Chapter 11.

**Exercise 1.0.4** (Inability to discretise nonabelian Lie groups)**.** Show that if $n \geq 3$, then the orthogonal group $\mathrm{O}_n(\mathbf{R})$ cannot contain arbitrarily dense finite subgroups, in the sense that there exists an $\varepsilon = \varepsilon_n > 0$ depending only on $n$ such that for every finite subgroup $G$ of $\mathrm{O}_n(\mathbf{R})$, there exists a ball of radius $\varepsilon$ in $\mathrm{O}_n(\mathbf{R})$ (with, say, the operator norm metric) that is disjoint from $G$. What happens in the $n = 2$ case?

**Remark 1.0.4.** More precise classifications of the finite subgroups of $\mathrm{U}_n(\mathbf{C})$ are known, particularly in low dimensions. For instance, it is a classical result that the only finite subgroups of $\mathrm{SO}_3(\mathbf{R})$ (which $\mathrm{SU}_2(\mathbf{C})$ is a double cover of) are isomorphic to either a cyclic group, a *dihedral group*, or the symmetry group of one of the *Platonic solids*.

## 1.1. Hilbert's fifth problem

One of the fundamental categories of objects in modern mathematics is the category of *Lie groups*, which are rich in both algebraic and analytic structure. Let us now briefly recall the precise definition of what a Lie group is.

**Definition 1.1.1** (Smooth manifold)**.** Let $d \geq 0$ be a natural number. A $d$-dimensional *topological manifold* is a *Hausdorff* topological space $M$ which

is *locally Euclidean*, thus every point in $M$ has a neighbourhood which is homeomorphic to an open subset of $\mathbf{R}^d$.

A *smooth atlas* on a $d$-dimensional topological manifold $M$ is a family $(\phi_\alpha)_{\alpha\in A}$ of homeomorphisms $\phi_\alpha : U_\alpha \to V_\alpha$ from open subsets $U_\alpha$ of $M$ to open subsets $V_\alpha$ of $\mathbf{R}^d$, such that the $U_\alpha$ form an open cover of $M$, and for any $\alpha, \beta \in A$, the map $\phi_\beta \circ \phi_\alpha^{-1}$ is smooth (i.e., infinitely differentiable) on the domain of definition $\phi_\alpha(U_\alpha \cap U_\beta)$. Two smooth atlases are *equivalent* if their union is also a smooth atlas; this is easily seen to be an equivalence relation. An equivalence class of smooth atlases is a *smooth structure*. A *smooth manifold* is a topological manifold equipped with a smooth structure.

A map $\psi : M \to M'$ from one smooth manifold to another is said to be smooth if $\phi'_\alpha \circ \psi \circ \phi_\beta^{-1}$ is a smooth function on the domain of definition $V_\beta \cap \phi_\beta^{-1}(U_\beta \cap \psi^{-1}(U_\alpha))$ for any smooth charts $\phi_\beta, \phi'_\alpha$ in any the smooth atlases of $M, M'$ respectively (one easily verifies that this definition is independent of the choice of smooth atlas in the smooth structure).

Note that we do not require manifolds to be connected, nor do we require them to be embeddable inside an ambient Euclidean space such as $\mathbf{R}^n$, although certainly many key examples of manifolds are of this form. The requirement that the manifold be Hausdorff is a technical one, in order to exclude pathological examples such as the line with a doubled point (formally, consider the double line $\mathbf{R} \times \{0,1\}$ after identifying $(x,0)$ with $(x,1)$ for all $x \in \mathbf{R}\backslash\{0\}$), which is locally Euclidean but not Hausdorff[1].

**Remark 1.1.2.** It is a plausible, but nontrivial, fact that a (nonempty) topological manifold can have at most one dimension $d$ associated to it; thus a manifold $M$ cannot both be locally homeomorphic to $\mathbf{R}^d$ and locally homeomorphic to $\mathbf{R}^{d'}$ unless $d = d'$. This fact is a consequence of Brouwer's *invariance of domain theorem*; see Exercise 6.0.4. On the other hand, it is an easy consequence of the *rank-nullity theorem* that a *smooth* manifold can have at most one dimension, without the need to invoke invariance of domain; we leave this as an exercise.

**Definition 1.1.3** (Lie group)**.** A *Lie group* is a group $G = (G, \cdot)$ which is also a smooth manifold, such that the group operations $\cdot : G \times G \to G$ and $()^{-1} : G \to G$ are smooth maps. (Note that the Cartesian product of two smooth manifolds can be given the structure of a smooth manifold in the obvious manner.) We will also use additive notation $G = (G, +)$ to describe some Lie groups, but only in the case when the Lie group is abelian.

[1] In some literature, additional technical assumptions such as *paracompactness*, *second countability*, or *metrisability* are imposed to remove pathological examples of topological manifolds such as the *long line*, but it will not be necessary to do so in this text, because (as we shall see later) we can essentially get such properties "for free" for locally Euclidean groups.

**Remark 1.1.4.** In some literature, Lie groups are required to be connected (and occasionally, are even required to be simply connected), but we will not adopt this convention here. One can also define infinite-dimensional Lie groups, but in this text all Lie groups are understood to be finite dimensional.

**Example 1.1.5.** Every group can be viewed as a Lie group if given the discrete topology (and the discrete smooth structure). (Note that we are not requiring Lie groups to be connected.)

**Example 1.1.6.** Finite-dimensional vector spaces such as $\mathbf{R}^d$ are (additive) Lie groups, as are sublattices such as $\mathbf{Z}^d$ or quotients such as $\mathbf{R}^d/\mathbf{Z}^d$. However, nonclosed subgroups such as $\mathbf{Q}^d$ are not manifolds (at least with the topology induced from $\mathbf{R}^d$) and are thus not Lie groups; similarly, quotients such as $\mathbf{R}^d/\mathbf{Q}^d$ are not Lie groups either (they are not even Hausdorff). Also, infinite-dimensional topological vector spaces (such as $\mathbf{R}^{\mathbf{N}}$ with the product topology) will not be Lie groups.

**Example 1.1.7.** The *general linear group* $\mathrm{GL}_n(\mathbf{C})$ of invertible $n \times n$ complex matrices is a Lie group. A theorem of Cartan (Theorem 3.0.14) asserts that any closed subgroup of a Lie group is a smooth submanifold of that Lie group and is in particular also a Lie group. In particular, closed linear groups (i.e., closed subgroups of a general linear group) are Lie groups; examples include the real general linear group $\mathrm{GL}_n(\mathbf{R})$, the *unitary group* $\mathrm{U}_n(\mathbf{C})$, the *special unitary group* $\mathrm{SU}_n(\mathbf{C})$, the *orthogonal group* $\mathrm{O}_n(\mathbf{R})$, the *special orthogonal group* $\mathrm{SO}_n(\mathbf{R})$, and the *Heisenberg group*

$$\begin{pmatrix} 1 & \mathbf{R} & \mathbf{R} \\ 0 & 1 & \mathbf{R} \\ 0 & 0 & 1 \end{pmatrix}$$

of unipotent upper triangular $3 \times 3$ real matrices. Many Lie groups are isomorphic to closed linear groups; for instance, the additive group $\mathbf{R}$ can be identified with the closed linear group

$$\begin{pmatrix} 1 & \mathbf{R} \\ 0 & 1 \end{pmatrix},$$

the circle $\mathbf{R}/\mathbf{Z}$ can be identified with $\mathrm{SO}_2(\mathbf{R})$ (or $\mathrm{U}_1(\mathbf{C})$), and so forth. However, not all Lie groups are isomorphic to closed linear groups. A somewhat trivial example is that of a discrete group with cardinality larger than the continuum, which is simply too large to fit inside any linear group. A less pathological example is provided by the Weil-Heisenberg group

$$G := \begin{pmatrix} 1 & \mathbf{R} & \mathbf{R}/\mathbf{Z} \\ 0 & 1 & \mathbf{R} \\ 0 & 0 & 1 \end{pmatrix} := \begin{pmatrix} 1 & \mathbf{R} & \mathbf{R} \\ 0 & 1 & \mathbf{R} \\ 0 & 0 & 1 \end{pmatrix} / \begin{pmatrix} 1 & 0 & \mathbf{Z} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3}$$

which is isomorphic to the image of the Heisenberg group under the *Weil representation*, or equivalently the group of isometries of $L^2(\mathbf{R})$ generated by translations and modulations. Despite this, though, it is helpful to think of closed linear groups and Lie groups as being almost the same concept as a first approximation. For instance, one can show using *Ado's theorem* (Theorem 13.0.4) that every Lie group is *locally* isomorphic to a linear *local* group (a concept we will discuss in Section 2.1).

An important subclass of the closed linear groups are the *linear algebraic groups*, in which the group is also a real or complex *algebraic variety* (or at least an algebraically constructible set). All of the examples of closed linear groups given above are linear algebraic groups, although there exist closed linear groups that are not isomorphic to any algebraic group; see Proposition 21.0.4.

**Exercise 1.1.1** (Weil-Heisenberg group is not linear)**.** Show that there is no injective homomorphism $\rho : G \to \mathrm{GL}_n(\mathbf{C})$ from the Weil-Heisenberg group (1.3) to a general linear group $\mathrm{GL}_n(\mathbf{C})$ for any finite $n$. (*Hint:* The centre $[G, G]$ maps via $\rho$ to a circle subgroup of $\mathrm{GL}_n(\mathbf{C})$; diagonalise this subgroup and reduce to the case when the image of the centre consists of multiples of the identity. Now, use the fact that commutators in $\mathrm{GL}_n(\mathbf{C})$ have determinant one.) This fact was first observed by Birkhoff.

*Hilbert's fifth problem*, like many of Hilbert's problems, does not have a unique interpretation, but one of the most commonly accepted interpretations of the question posed by Hilbert is to determine if the requirement of smoothness in the definition of a Lie group is redundant. (There is also an analogue of Hilbert's fifth problem for group *actions*, known as the *Hilbert-Smith conjecture*; see Chapter 17.) To answer this question, we need to relax the notion of a Lie group to that of a *topological group*.

**Definition 1.1.8** (Topological group)**.** A *topological group* is a group $G = (G, \cdot)$ that is also a topological space, in such a way that the group operations $\cdot : G \times G \to G$ and $()^{-1} : G \to G$ are continuous. (As before, we also consider additive topological groups $G = (G, +)$ provided that they are abelian.)

Clearly, every Lie group is a topological group if one simply forgets the smooth structure, retaining only the topological and group structures. Furthermore, such topological groups remain locally Euclidean. It was established by Montgomery-Zippin [**MoZi1952**] and Gleason [**Gl1952**] that the converse statement holds, thus solving at least one formulation of Hilbert's fifth problem:

**Theorem 1.1.9** (Hilbert's fifth problem)**.** *Let $G$ be an object with the following properties:*

(1) *(Group-like object) $G$ is a topological group.*

(2) *(Weak regularity) $G$ is locally Euclidean.*

*Then*

(i) *(Lie-type structure) $G$ is isomorphic to a Lie group.*

**Exercise 1.1.2.** Show that a locally Euclidean topological group is necessarily Hausdorff (without invoking Theorem 1.1.9).

We will prove this theorem in Section 6. As it turns out, Theorem 1.1.9 is not directly useful for many applications, because it is often difficult to verify that a given topological group is locally Euclidean. On the other hand, the weaker property of *local compactness*, which is clearly implied by the locally Euclidean property, is much easier to verify in practice. One can then ask the more general question of whether every locally compact group is isomorphic to a Lie group. Unfortunately, the answer to this question is easily seen to be no, as the following examples show:

**Example 1.1.10** (Trivial topology)**.** A group equipped with the trivial topology is a compact (hence locally compact) group, but will not be Hausdorff (and thus not Lie) unless the group is also trivial. Of course, this is a rather degenerate counterexample and can be easily eliminated in practice. For instance, we will see later that any topological group can be made Hausdorff by quotienting out the closure of the identity.

**Example 1.1.11** (Infinite-dimensional torus)**.** The infinite-dimensional torus $(\mathbf{R}/\mathbf{Z})^{\mathbf{N}}$ (with the product topology) is an (additive) topological group, which is compact (and thus locally compact) by *Tychonoff's theorem*. However, it is not a Lie group.

**Example 1.1.12** ($p$-adics)**.** Let $p$ be a prime. We define the *$p$-adic norm* $\|\|_p$ on the integers $\mathbf{Z}$ by defining $\|n\|_p := p^{-j}$, where $p^j$ is the largest power of $p$ that divides $n$ (with the convention $\|0\|_p := 0$). This is easily verified to generate a metric (and even an *ultrametric*) on $\mathbf{Z}$; the $p$-adic integers $\mathbf{Z}_p$ are then defined as the *metric completion* of $\mathbf{Z}$ under this metric. This is easily seen to be a compact (hence locally compact) additive group (topologically, it is homeomorphic to a *Cantor set*). However, it is not locally Euclidean (or even *locally connected*), and so is not isomorphic to a Lie group.

One can also extend the $p$-adic norm to the ring $\mathbf{Z}[\frac{1}{p}]$ of rationals of the form $a/p^j$ for some integers $a, j$ in the obvious manner; the metric completion of this space is then the $p$-adic rationals $\mathbf{Q}_p$. This is now a locally compact additive group rather than a compact one ($\mathbf{Z}_p$ is a compact open neighbourhood of the identity); it is still not locally connected, so it is still not a Lie group.

One can also define algebraic groups such as $\mathrm{GL}_n$ over the $p$-adic rationals $\mathbf{Q}_p$; thus for instance $\mathrm{GL}_n(\mathbf{Q}_p)$ is the group of invertible $n \times n$ matrices with entries in the $p$-adics. This is still a locally compact group, and is certainly not Lie.

**Exercise 1.1.3** (Solenoid)**.** Let $p$ be a prime. Let $G$ be the *solenoid group* $G := (\mathbf{Z}_p \times \mathbf{R})/\mathbf{Z}^\Delta$, where $\mathbf{Z}^\Delta := \{(n,n) : n \in \mathbf{Z}\}$ is the diagonally embedded copy of the integers in $\mathbf{Z}_p \times \mathbf{R}$. (Topologically, $G$ can be viewed as the set $\mathbf{Z}_p \times [0,1]$ after identifying $(x+1,1)$ with $(x,0)$ for all $x \in \mathbf{Z}_p$.) Show that $G$ is a compact additive group that is connected but not locally connected (and thus not a Lie group). Thus one cannot eliminate $p$-adic type behaviour from locally compact groups simply by restricting attention to the connected case (although we will see later that one can do so by restricting to the *locally* connected case).

We have now seen several examples of locally compact groups that are not Lie groups. However, all of these examples are "almost" Lie groups in that they can be turned into Lie groups by quotienting out a small compact normal subgroup. (It is easy to see that the quotient of a locally compact group by a compact normal subgroup is again a locally compact group.) For instance, a group with the trivial topology becomes Lie after quotienting out the entire group (which is "small" in the sense that it is contained in every open neighbourhood of the origin). The infinite-dimensional torus $(\mathbf{R}/\mathbf{Z})^{\mathbf{N}}$ can be quotiented into a finite-dimensional torus $(\mathbf{R}/\mathbf{Z})^d$ (which is of course a Lie group) by quotienting out the compact subgroup $\{0\}^d \times (\mathbf{R}/\mathbf{Z})^{\mathbf{N}}$; note from the definition of the product topology that these compact subgroups shrink to zero in the sense that every neighbourhood of the group identity contains at least one (and in fact all but finitely many) of these subgroups. Similarly, with the $p$-adic group $\mathbf{Z}_p$, one can quotient out by the compact (and open) subgroups $p^j\mathbf{Z}_p$ (which also shrink to zero, as discussed above) to obtain the cyclic groups $\mathbf{Z}/p^j\mathbf{Z}$, which are discrete and thus Lie. Quotienting out $\mathbf{Q}_p$ by the same compact open subgroups $p^j\mathbf{Z}_p$ also leads to discrete (hence Lie) quotients; similarly for algebraic groups defined over $\mathbf{Q}_p$, such as $\mathrm{GL}_n(\mathbf{Q}_p)$. Finally, with the solenoid group $G := (\mathbf{Z}_p \times \mathbf{R})/\mathbf{Z}^\Delta$, one can quotient out the copy of $p^j\mathbf{Z}_p \times \{0\}$ in $G$ for $j = 0, 1, 2, \ldots$ (which are another sequence of compact subgroups shrinking to zero) to obtain the quotient group $(\mathbf{Z}/p^j\mathbf{Z} \times \mathbf{R})/\mathbf{Z}^\Delta$, which is isomorphic to a (highly twisted) circle $\mathbf{R}/\mathbf{Z}$ and is thus Lie.

Inspired by these examples, we might be led to the following conjecture: if $G$ is a locally compact group, and $U$ is a neighbourhood of the identity, then there exists a compact normal subgroup $K$ of $G$ contained in $U$ such that $G/K$ is a Lie group. In the event that $G$ is Hausdorff, this is equivalent to asserting that $G$ is the *projective limit* (or *inverse limit*) of Lie groups.

This conjecture is true in several cases; for instance, one can show using the *Peter-Weyl theorem* (which we will discuss in Chapter 4) that it is true for compact groups, and we will later see that it is also true for connected locally compact groups (see Theorem 6.0.11). However, it is not quite true in general, as the following example shows.

**Exercise 1.1.4.** Let $p$ be a prime, and let $T : \mathbf{Q}_p \to \mathbf{Q}_p$ be the automorphism $Tx := px$. Let $G := \mathbf{Q}_p \rtimes_T \mathbf{Z}$ be the semidirect product of $\mathbf{Q}_p$ and $\mathbf{Z}$ twisted by $T$; more precisely, $G$ is the Cartesian product $\mathbf{Q}_p \times \mathbf{Z}$ with the product topology and the group law

$$(x, n)(y, m) := (x + T^n y, n + m).$$

Show that $G$ is a locally compact group which is not isomorphic to a Lie group, and that $\mathbf{Z}_p \times \{0\}$ is an open neighbourhood of the identity that contains no nontrivial normal subgroups of $G$. Conclude that the conjecture stated above is false.

The difficulty in the above example was that it was not easy to keep a subgroup normal with respect to the entire group $\mathbf{Q}_p \rtimes_T \mathbf{Z}$. Note however that $G$ contains a "large" (and more precisely, *open*) subgroup $\mathbf{Q}_p \times \{0\}$ which *is* the projective limit of Lie groups. So the above examples do not rule out that the conjecture can still be salvaged if one passes from a group $G$ to an open subgroup $G'$. This is indeed the case:

**Theorem 1.1.13** (Gleason-Yamabe theorem [**Gl1951, Ya1953b**]). *Let $G$ obey the following hypotheses:*

(1) *(Group-like object) $G$ is a topological group.*

(2) *(Weak regularity) $G$ is locally compact.*

*Then for every open neighbourhood $U$ of the identity, there exists a subgroup $G'$ of $G$ and a compact normal subgroup $K$ of $G'$ with the following properties:*

(i) *($G'/K$ is close to $G$) $G'$ is an open subgroup of $G$, and $K$ is contained in $U$.*

(ii) *(Lie-type structure) $G'/K$ is isomorphic to a Lie group.*

The proof of this theorem will occupy the next few sections of this text, being finally proven in Chapter 5. As stated, $G'$ may depend on $U$, but one can in fact take the open subgroup $G'$ to be uniform in the choice of $U$; we will show this in later sections. Theorem 1.1.9 can in fact be deduced from Theorem 1.1.13 and some topological arguments involving the invariance of domain theorem; this will be shown in Chapter 6.

The Gleason-Yamabe theorem asserts that locally compact groups are "essentially" Lie groups, after ignoring the very large scales (by restricting

to an open subgroup) and also ignoring the very small scales (by allowing one to quotient out by a small group). In special cases, the conclusion of the theorem can be simplified. For instance, it is easy to see that an open subgroup $G'$ of a topological group $G$ is also closed (since the complement $G \backslash G'$ is a union of cosets of $G'$), and so if $G$ is connected, there are no open subgroups other than $G$ itself. Thus, in the connected case of Theorem 1.1.13, one can take $G = G'$. In a similar spirit, if $G$ has the *no small subgroups* (NSS) property, that is to say that there exists an open neighbourhood of the identity that contains no nontrivial subgroups of $G$, then we can take $K$ to be trivial. Thus, as a special case of the Gleason-Yamabe theorem, we see that all connected NSS locally compact groups are Lie; in fact it is not difficult to then conclude that any locally compact NSS group (regardless of connectedness) is Lie. Conversely, this claim (which we isolate as Corollary 5.3.3) turns out to be a key step in the *proof* of Theorem 1.1.13, as we shall see later. (It is also not difficult to show that all Lie groups are NSS; see Exercise 5.3.1.)

The proof of the Gleason-Yamabe theorem proceeds in a somewhat lengthy series of steps in which the initial regularity (local compactness) on the group $G$ is gradually upgraded to increasingly stronger regularity (e.g. metrisability, the NSS property, or the locally Euclidean property) until one eventually obtains Lie structure; see Figure 1. A key turning point in the argument will be the construction of a metric (which we call a *Gleason metric*) on (a large portion of) $G$ which obeys a commutator estimate similar to (1.2).

While the Gleason-Yamabe theorem does not completely classify all locally compact groups (as mentioned earlier, it primarily controls the medium-scale behaviour, and not the very fine-scale or very coarse-scale behaviour, of such groups), it is still powerful enough for a number of applications, to which we now turn.

## 1.2. Approximate groups

We now discuss what appears at first glance to be an unrelated topic, namely that of *additive combinatorics* (and its noncommutative counterpart, multiplicative combinatorics). One of the main objects of study in either additive or multiplicative combinatorics are *approximate groups* — sets $A$ (typically finite) contained in an additive or multiplicative ambient group $G$ that are "almost groups" in the sense that they are "almost" closed under either addition or multiplication. (One can also consider abstract approximate groups that are not contained in an ambient genuine group, but we will not do so here.)

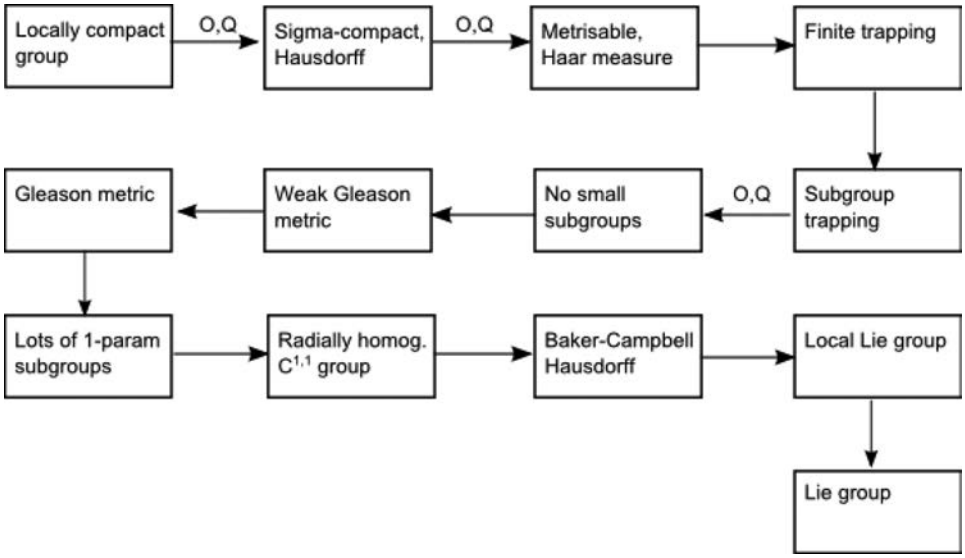


**Figure 1.** A schematic description of the steps needed to establish the Gleason-Yamabe theorem. The annotation $O, Q$ on an arrow indicates that one has to pass to an open subgroup, and then quotient out a compact normal subgroup, in order to obtain the additional structure at the end of the arrow.

There are several ways to quantify what it means for a set $A$ to be "almost" closed under addition or multiplication. Here are some common formulations of this idea (phrased in multiplicative notation, for the sake of concreteness):

(1) (Statistical multiplicative structure) For a "large" proportion of pairs $(a, b) \in A \times A$, the product $ab$ also lies in $A$.

(2) (Small product set) The "size" of the product set $A \cdot A := \{ab : a, b \in A\}$ is "comparable" to the "size" of the original set $A$. (For technical reasons, one sometimes uses the triple product $A \cdot A \cdot A := \{abc : a, b, c \in A\}$ instead of the double product.)

(3) (Covering property) The product set $A \cdot A$ can be covered by a "bounded" number of (left or right) translates of the original set $A$.

Of course, to make these notions precise one would have to precisely quantify the various terms in quotes. Fortunately, the basic theory of additive combinatorics (and multiplicative combinatorics) can be used to show that all these different notions of additive or multiplicative structure are "essentially" equivalent; see [**TaVu2006**, Chapter 2] or [**Ta2008b**] for more discussion.

For the purposes of this text, it will be convenient to focus on the use of covering to describe approximate multiplicative structure. More precisely:

**Definition 1.2.1** (Approximate groups)**.** Let $G$ be a multiplicative group, and let $K \geq 1$ be a real number. A *$K$-approximate subgroup* of $G$, or *$K$-approximate group* for short, is a subset $A$ of $G$ which contains the identity, is symmetric (thus $A^{-1} := \{a^{-1} : a \in A\}$ is equal to $A$) and is such that $A \cdot A$ can be covered by at most $K$ left-translates (or equivalently by symmetry, right translates) of $A$, thus there exists a subset $X$ of $G$ of cardinality at most $K$ such that $A \cdot A \subset X \cdot A$.

In most combinatorial applications, one only considers approximate groups that are finite sets, but one could certainly also consider countably or uncountably infinite approximate groups. We remark that this definition is essentially from [**Ta2008b**] (although the definition in [**Ta2008b**] places some additional minor constraints on the set $X$ which have turned out not to be terribly important in practice).

**Example 1.2.2.** A 1-approximate subgroup of $G$ is the same thing as a genuine subgroup of $G$.

**Example 1.2.3.** In the additive group of the integers $\mathbf{Z}$, the symmetric arithmetic progression $\{-N, \ldots, N\}$ is a 2-approximate group for any $N \geq 1$. More generally, in any additive group $G$, the *symmetric generalised arithmetic progression*

$$\{a_1 v_1 + \cdots + a_r v_r : a_1, \ldots, a_r \in \mathbf{Z}, |a_i| \leq N_i \forall i = 1, \ldots, r\}$$

with $v_1, \ldots, v_r \in G$ and $N_1, \ldots, N_r > 0$, is a $2^r$-approximate group.

**Exercise 1.2.1.** Let $A$ be a convex symmetric subset of $\mathbf{R}^d$. Show that $A$ is a $5^d$-approximate group. (*Hint:* Greedily pack $2A$ with disjoint translates of $\frac{1}{2}A$.)

**Example 1.2.4.** If $A$ is an open *precompact*[2] symmetric neighbourhood of the identity in a locally compact group $G$, then $A$ is a $K$-approximate group for some finite $K$. Thus we see some connection between locally compact groups and approximate groups; we will see a deeper connection involving *ultraproducts* in Chapter 7.

**Example 1.2.5.** Let $G$ be a $d$-dimensional Lie group. Then $G$ is a smooth manifold, and can thus be (nonuniquely) given the structure of a *Riemannian manifold*. If one does so, then for sufficiently small radii $r$, the ball $B(1, r)$ around the identity 1 will be a $O_d(1)$-approximate group.

[2] A subset of a topological space is said to be precompact if its closure is compact.

**Example 1.2.6** (Extensions)**.** Let $\phi : G \to H$ be a surjective group homomorphism (thus $G$ is a *group extension* of $H$ by the kernel $\ker(\phi)$ of $\phi$). If $A$ is a $K$-approximate subgroup of $H$, then $\phi^{-1}(A)$ is a $K$-approximate subgroup of $G$. One can think of $\phi^{-1}(A)$ as an extension of the approximate group $A$ by $\ker(\phi)$.

The classification of approximate groups is of importance in additive combinatorics, and has connections with number theory, geometric group theory, and the theory of expander graphs. One can ask for a quantitative classification, in which one has explicit dependence of constants on the approximate group parameter $K$, or one can settle for a qualitative classification in which one does not attempt to control this dependence of constants. In this text we will focus on the latter question, as this allows us to bring in qualitative tools such as the Gleason-Yamabe theorem to bear on the problem.

In the abelian case when the ambient group $G$ is additive, approximate groups are classified by *Freiman's theorem for abelian groups*[3] [**GrRu2007**]. As before, we phrase this theorem in a slightly stilted fashion (and in a qualitative, rather than quantitative, manner) in order to demonstrate its alignment with the general principles stated in the introduction.

**Theorem 1.2.7** (Freiman's theorem in an abelian group)**.** *Let $A$ be an object with the following properties:*

(1) *(Group-like object) $A$ is a subset of an additive group $G$.*

(2) *(Weak regularity) $A$ is a $K$-approximate group.*

(3) *(Discreteness) $A$ is finite.*

*Then there exists a finite subgroup $H$ of $G$, and a subset $P$ of $G/H$, with the following properties:*

(i) *(P is close to A/H) $\pi^{-1}(P)$ is contained in $4A := A + A + A + A$, where $\pi : G \to G/H$ is the quotient map, and $|P| \gg_K |A|/|H|$.*

(ii) *(Nilpotent type structure) $P$ is a symmetric generalised arithmetic progression of rank $O_K(1)$ (see Example 1.2.3).*

Informally, this theorem asserts that in the abelian setting, discrete approximate groups are essentially bounded rank symmetric generalised arithmetic progressions, extended by finite groups (such extensions are also known as *coset progressions*). The theorem has a simpler conclusion (and is simpler to prove) in the case when $G$ is a torsion-free abelian group (such as $\mathbf{Z}$), since in this case $H$ is trivial.

[3]The original theorem of Freiman [**Fr1973**] obtained an analogous classification in the case when $G$ was a torsion-free abelian group, such as the integers $\mathbf{Z}$.

We will not discuss the proof of Theorem 1.2.7 from [**GrRu2007**] here, save to say that it relies heavily on Fourier-analytic methods, and as such, does not seem to easily extend to a general nonabelian setting. To state the nonabelian analogue of Theorem 1.2.7, one needs multiplicative analogues of the concept of a generalised arithmetic progression. An ordinary (symmetric) arithmetic progression $\{-Nv, \dots, Nv\}$ has an obvious multiplicative analogue, namely a (symmetric) geometric progression $\{a^{-N}, \dots, a^N\}$ for some generator $a \in G$. In a similar vein, if one has $r$ *commuting* generators $a_1, \dots, a_r$ and some dimensions $N_1, \dots, N_r > 0$, one can form a symmetric generalised geometric progression

$$P := \{a_1^{n_1} \dots a_r^{n_r} : |n_i| \leq N_i \forall 1 \leq i \leq r\}, \tag{1.4}$$

which will still be a $3^r$-approximate group. However, if the $a_1, \dots, a_r$ do not commute, then the set $P$ defined in (1.4) is not quite the right concept to use here; for instance, there it is no reason for $P$ to be symmetric. However, it can be modified as follows:

**Definition 1.2.8** (noncommutative progression)**.** Let $a_1, \dots, a_r$ be elements of a (not necessarily abelian) group $G = (G, \cdot)$, and let $N_1, \dots, N_r > 0$. We define the *noncommutative progression* $P = P(a_1, \dots, a_r; N_1, \dots, N_r)$ of rank $r$ with generators $a_1, \dots, a_r$ and dimensions $N_1, \dots, N_r$ to be the collection of all words $w$ composed using the alphabet $a_1, a_1^{-1}, \dots, a_r, a_r^{-1}$, such that for each $1 \leq i \leq r$, the total number of occurrences of $a_i$ and $a_i^{-1}$ combined in $w$ is at most $N_i$.

**Example 1.2.9.** $P(a, b; 1, 2)$ consists of the elements

$$1, a^{\pm}, b^{\pm}, b^{\pm}a^{\pm}, a^{\pm}b^{\pm}, b^{\pm 2}, b^{\pm 2}a^{\pm}, b^{\pm}a^{\pm}b^{\pm}, a^{\pm}b^{\pm 2},$$

where each occurrence of $\pm$ can independently be set to $+$ or $-$; thus $P(a, b; 1, 2)$ can have as many as 31 elements.

**Example 1.2.10.** If the $a_1, \dots, a_r$ commute, then the noncommutative progression $P(a_1, \dots, a_r; N_1, \dots, N_r)$ simplifies to (1.4).

**Exercise 1.2.2.** Let $G$ be the discrete Heisenberg group

$$G = \begin{pmatrix} 1 & \mathbf{Z} & \mathbf{Z} \\ 0 & 1 & \mathbf{Z} \\ 0 & 0 & 1 \end{pmatrix} \tag{1.5}$$

and let

$$e_1 := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, e_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

be the two generators of $G$. Let $N \geq 1$ be a sufficiently large natural number. Show that the noncommutative progression $P(e_1, e_2; N, N)$ contains all the

group elements $\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$ of $G$ with $|a|, |b| \leq \delta N$ and $|c| \leq \delta N^2$ for a sufficiently small absolute constant $\delta > 0$; conversely, show that all elements of $P(e_1, e_2; N, N)$ are of the form $\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$ with $|a|, |b| \leq CN$ and $|c| \leq CN^2$ for some sufficiently large absolute constant $C > 0$. Thus, informally, we have

$$P(e_1, e_2; N, N) = \begin{pmatrix} 1 & O(N) & O(N^2) \\ 0 & 1 & O(N) \\ 0 & 0 & 1 \end{pmatrix}.$$

It is clear that noncommutative progressions $P(a_1, \ldots, a_r; N_1, \ldots, N_r)$ are symmetric and contain the identity. However, if the $a_1, \ldots, a_r$ do not have any commutative properties, then the size of these progressions can grow exponentially in $N_1, \ldots, N_r$ and will not be approximate groups with any reasonable parameter $K$. However, the situation changes when the $a_1, \ldots, a_r$ generate a *nilpotent group*[4]:

**Proposition 1.2.11.** *Suppose that $a_1, \ldots, a_r \in G$ generate a nilpotent group of step $s$, and suppose that $N_1, \ldots, N_r$ are all sufficiently large depending on $r, s$. Then $P(a_1, \ldots, a_r; N_1, \ldots, N_r)$ is an $O_{r,s}(1)$-approximate group.*

We will prove this proposition in Chapter 12.

We can now state the noncommutative analogue of Theorem 1.2.7, proven in [**BrGrTa2011**]:

**Theorem 1.2.12** (Freiman's theorem in an arbitrary group)**.** *Let $A$ be an object with the following properties:*

(1) *(Group-like object) $A$ is a subset of a multiplicative group $G$.*

(2) *(Weak regularity) $A$ is a $K$-approximate group.*

(3) *(Discreteness) $A$ is finite.*

*Then there exists a finite subgroup $H$ of $G$, and a subset $P$ of $N(H)/H$ (where $N(H) := \{g \in G : gH = Hg\}$ is the normaliser of $H$), with the following properties:*

(i) *(P is close to A/H) $\pi^{-1}(P)$ is contained in $A^4 := A \cdot A \cdot A \cdot A$, where $\pi : N(H) \to N(H)/H$ is the quotient map, and $|P| \gg_K |A|/|H|$.*

(ii) *(Nilpotent type structure) $P$ is a noncommutative progression of rank $O_K(1)$, whose generators generate a nilpotent group of step $O_K(1)$.*

[4] A group $G$ is nilpotent if the lower central series $G_1 := G, G_2 := [G, G_1], G_3 := [G, G_2], \ldots$, etc. eventually becomes trivial.

The proof of this theorem relies on the Gleason-Yamabe theorem (Theorem 1.1.13), and will be discussed in Chapter 8. The key connection will take some time to explain properly, but roughly speaking, it comes from the fact that the *ultraproduct* of a sequence of $K$-approximate groups can be used to generate a locally compact group, to which the Gleason-Yamabe theorem can be applied. This in turn can be used to place a metric on approximate groups that obeys a commutator estimate similar to (1.2), which allows one to run an argument similar to that used to prove Theorem 1.0.2.

## 1.3. Gromov's theorem

The final topic of this chapter will be *Gromov's theorem on groups of polynomial growth* [**Gr1981**]. This theorem is analogous to Theorem 1.1.13 or Theorem 1.2.12, but in the category of *finitely generated groups* rather than locally compact groups or approximate groups.

Let $G$ be a group that is generated by a finite set $S$ of generators; for notational simplicity we will assume that $S$ is symmetric and contains the origin. Then $S$ defines a (right-invariant) *word metric* on $G$, defined by setting $d(x,y)$ for $x,y \in G$ to be the least natural number $n$ such that $x \in S^n y$. One easily verifies that this is indeed a metric that is right-invariant (thus $d(xg,yg) = d(x,y)$ for all $x,y,g \in G$). Geometrically, this metric describes the geometry of the *Cayley graph* on $G$ formed by connecting $x$ to $sx$ for each $x \in G$ and $s \in S$. (See [**Ta2011c**, §2.3] for more discussion of using Cayley graphs to study groups geometrically.)

Let us now consider the growth of the balls $B(1,R) = S^{\lfloor R \rfloor}$ as $R \to \infty$, where $\lfloor R \rfloor$ is the integer part of $R$. On the one hand, we have the trivial upper bound

$$|B(1,R)| \leq |S|^R$$

that shows that such balls can grow at most exponentially. And for "typical" nonabelian groups, this exponential growth actually occurs; consider the case for instance when $S$ consists of the generators of a free group (together with their inverses, and the group identity). However, there are some groups for which the balls grow at a much slower rate. A somewhat trivial example is that of a finite group $G$, since clearly $|B(1,R)|$ will top out at $|G|$ (when $R$ reaches the *diameter* of the Cayley graph) and stop growing after that point. Another key example is the abelian case:

**Exercise 1.3.1.** If $G$ is an abelian group generated by a finite symmetric set $S$ containing the identity, show that

$$|B(1,R)| \leq (1+R)^{|S|}.$$

In particular, $B(1,R)$ grows at a polynomial rate in $R$.

Let us say that a finite group $G$ is a *group of polynomial growth* if one has $|B(1,R)| \leq CR^d$ for all $R \geq 1$ and some constants $C, d > 0$.

**Exercise 1.3.2.** Show that the notion of a group of polynomial growth (as well as the rate $d$ of growth) does not depend on the choice of generators $S$; thus if $S'$ is another set of generators for $G$, show that $G$ has polynomial growth with respect to $S$ with rate $d$ if and only if it has polynomial growth with respect to $S'$ with rate $d$.

**Exercise 1.3.3.** Let $G$ be a finitely generated group, and let $G'$ be a finite index subgroup of $G$.

(i) Show that $G'$ is also finitely generated. (*Hint:* Let $S$ be a symmetric set of generators for $G$ containing the identity, and locate a finite integer $n$ such that $S^{n+1}G' = S^nG'$. Then show that the set $S' := G' \cap S^{2n+1}$ is such that $S^{n+1} \subset S^nS'$. Conclude that $S^n$ meets every coset of $\langle S' \rangle$ (or equivalently that $G = S^n\langle S' \rangle$), and use this to show that $S'$ generates $G'$.)

(ii) Show that $G$ has polynomial growth if and only if $G'$ has polynomial growth.

(iii) More generally, show that any finitely generated subgroup of a group of polynomial growth also has polynomial growth. Conclude in particular that a group of polynomial growth cannot contain the free group on two generators.

From Exercise 1.2.2 we see that the discrete Heisenberg group (1.5) is of polynomial growth. It is in fact not difficult to show that, more generally, any nilpotent finitely generated group is of polynomial growth. By Exercise 1.3.3, this implies that any *virtually nilpotent* finitely generated group is of polynomial growth.

Gromov's theorem asserts the converse statement:

**Theorem 1.3.1** (Gromov's theorem [**Gr1981**])**.** *Let $G$ be an object with the following properties:*

(1) *(Group-like object) $G$ is a finitely generated group.*

(2) *(Weak regularity) $G$ is of polynomial growth.*

*Then there exists a subgroup $G'$ of $G$ such that*

(i) *($G'$ is close to $G$) The index $|G/G'|$ is finite.*

(ii) *(Nilpotent type structure) $G'$ is nilpotent.*

More succinctly: A finitely generated group is of polynomial growth if and only if it is virtually nilpotent.

Groups of polynomial growth are related to approximate groups by the following observation.

**Exercise 1.3.4** (Pigeonhole principle)**.** Let $G$ be a finitely generated group of polynomial growth, and let $S$ be a symmetric set of generators for $G$ containing the identity.

(i) Show that there exists a $C>1$ such that $|B(1,5R/2)| \leq C|B(1,R/2)|$ for a sequence $R = R_n$ of radii going to infinity.

(ii) Show that there exists a $K>1$ such that $B(1,R)$ is a $K$-approximate group for a sequence $R = R_n$ of radii going to infinity. (*Hint:* Argue as in Exercise 1.2.1.)

In Chapter 9 we will use this connection to deduce Theorem 1.3.1 from Theorem 1.2.12. From a historical perspective, this was not the first proof of Gromov's theorem; Gromov's original proof in [**Gr1981**] relied instead on a variant of Theorem 1.1.9 (as did some subsequent variants of Gromov's argument, such as the nonstandard analysis variant in [**vdDrWi1984**]), and a subsequent proof of Kleiner [**Kl2010**] went by a rather different route, based on earlier work of Colding and Minicozzi [**CoMi1997**] on harmonic functions of polynomial growth. (This latter proof is discussed in [**Ta2009**, §1.2] and [**Ta2011c**, §2.5].) The proof we will give in this text is more recent, based on an argument of Hrushovski [**Hr2012**]. We remark that the strategy used to prove Theorem 1.2.12 — namely taking an ultralimit of a sequence of approximate groups — also appears in Gromov's original argument[5]. We will discuss these sorts of limits more carefully in Chapter 7, but an informal example to keep in mind for now is the following: If one takes a discrete group (such as $\mathbf{Z}^d$) and rescales it (say to $\frac{1}{N}\mathbf{Z}^d$ for a large parameter $N$), then intuitively this rescaled group "converges" to a continuous group (in this case $\mathbf{R}^d$). More generally, one can generate locally compact groups (or at least locally compact spaces) out of the limits of (suitably normalised) groups of polynomial growth or approximate groups, which is one of the basic observations that tie the three different topics discussed above together.

As we shall see in Chapter 10, finitely generated groups arise naturally as the *fundamental groups* of compact manifolds. Using the tools of Riemannian geometry (such as the *Bishop-Gromov inequality*), one can relate the growth of such groups to the curvature of a metric on such a manifold. As a consequence, Gromov's theorem and its variants can lead to some nontrivial conclusions about the relationship between the topology of a manifold and its geometry. The following simple consequence is typical:

[5]Strictly speaking, he uses *Gromov-Hausdorff limits* instead of ultralimits, but the two types of limits are closely related, as we shall see in Chapter 7.

**Proposition 1.3.2.** *Let $M$ be a compact Riemannian manifold of nonnegative* Ricci curvature*. Then the fundamental group $\pi_1(M)$ of $M$ is virtually nilpotent.*

We will discuss this result and some related results (such as a relaxation of the nonnegative curvature hypothesis to an almost nonnegative curvature hypothesis) in Section 10. We also remark that the above proposition can also be proven (with stronger conclusions) by more geometric means, but there are some results of the above type which currently have no known proof that does not employ some version of Gromov's theorem at some point.

*Chapter 2*

# Lie groups, Lie algebras, and the Baker-Campbell-Hausdorff formula

In this chapter, we describe the basic analytic structure theory of Lie groups, by relating them to the simpler concept of a *Lie algebra*. Roughly speaking, the Lie algebra encodes the "infinitesimal" structure of a Lie group, but is a simpler object, being a vector space rather than a nonlinear manifold. Nevertheless, thanks to the fundamental theorems of Lie, the Lie algebra can be used to reconstruct the Lie group (at a local level, at least), by means of the *exponential map* and the *Baker-Campbell-Hausdorff formula*. As such, the local theory of Lie groups is completely described (in principle, at least) by the theory of Lie algebras, which leads to a number of useful consequences, such as the following:

(1) (Local Lie implies Lie) A topological group $G$ is Lie (i.e., it is isomorphic to a Lie group) if and only if it is locally Lie (i.e., the group operations are smooth near the origin).

(2) (Uniqueness of Lie structure) A topological group has at most one smooth structure on it that makes it Lie.

(3) (Weak regularity implies strong regularity, I) Lie groups are automatically real analytic. (In fact one only needs a "local $C^{1,1}$" regularity on the group structure to obtain real analyticity.)

(4) (Weak regularity implies strong regularity, II) A continuous homomorphism from one Lie group to another is automatically smooth (and real analytic).

The connection between Lie groups and Lie algebras also highlights the role of *one-parameter subgroups* of a topological group, which will play a central role in the solution of Hilbert's fifth problem (cf. Figure 1).

**Remark 2.0.3.** There is also a very important *algebraic* structure theory of Lie groups and Lie algebras, in which the Lie algebra is split into *solvable* and *semisimple* components, with the latter being decomposed further into *simple components*, which can then be completely classified using *Dynkin diagrams*. This classification is of fundamental importance in many areas of mathematics (e.g., representation theory, arithmetic geometry, and group theory), and many of the deeper facts about Lie groups and Lie algebras are proven via this classification (although in such cases it can be of interest to also find alternate proofs that avoid the classification). However, it turns out that we will not need this theory here, and so we will not discuss it further (though it can of course be found in any graduate text on Lie groups and Lie algebras, e.g., [**Bo1968**]).

## 2.1. Local groups

The connection between Lie groups and Lie algebras will be *local* in nature — the only portion of the Lie group that will be of importance will be the portion that is close to the group identity 1. To formalise this locality, it is convenient to introduce the notion of a *local group* and a *local Lie group*, which are local versions of the concept of a topological group and a Lie group respectively. We will only set up the barest bones of the theory of local groups here; a more detailed discussion is given in Chapter 15.

**Definition 2.1.1** (Local group)**.** A *local topological group*

$$G = (G, \Omega, \Lambda, 1, \cdot, ()^{-1}),$$

or *local group* for short, is a topological space $G$ equipped with an identity element $1 \in G$, a partially defined but continuous multiplication operation $\cdot : \Omega \to G$ for some domain $\Omega \subset G \times G$, and a partially defined but continuous inversion operation $()^{-1} : \Lambda \to G$, where $\Lambda \subset G$, obeying the following axioms:

(1) (Local closure) $\Omega$ is an open neighbourhood of $G \times \{1\} \cup \{1\} \times G$, and $\Lambda$ is an open neighbourhood of 1.

(2) (Local associativity) If $g, h, k \in G$ are such that $(g \cdot h) \cdot k$ and $g \cdot (h \cdot k)$ are both well-defined in $G$, then they are equal. (Note however that

it may be possible for one of these products to be defined but not the other.)

(3) (Identity) For all $g \in G$, $g \cdot 1 = 1 \cdot g = g$.

(4) (Local inverse) If $g \in G$ and $g^{-1}$ are well-defined in $G$, then[1] $g \cdot g^{-1} = g^{-1} \cdot g = 1$. (In particular this, together with the other axioms, forces $1^{-1} = 1$.)

We will sometimes use additive notation for local groups if the groups are abelian (by which we mean the statement that if $g + h$ is defined, then $h + g$ is also defined and equal to $g + h$.)

A local group is said to be *symmetric* if $\Lambda = G$, i.e., if every element $g$ in $G$ has an inverse $g^{-1}$ that is also in $G$.

A *local Lie group* is a local group that is also a smooth manifold, in such a fashion that the partially defined group operations $\cdot, ()^{-1}$ are smooth on their domain of definition.

Clearly, every topological group is a local group, and every Lie group is a local Lie group. We will sometimes refer to the former concepts as *global* topological groups and *global* Lie groups in order to distinguish them from their local counterparts. One could also consider local discrete groups, in which the topological structure is just the discrete topology, but we will not need to study such objects in here.

A model class of examples of a local (Lie) group comes from *restricting* a global (Lie) group to an open neighbourhood of the identity. Let us formalise this concept:

**Definition 2.1.2** (Restriction)**.** If $G$ is a local group, and $U$ is an open neighbourhood of the identity in $G$, then we define the *restriction* $G \restriction_U$ of $G$ to $U$ to be the topological space $U$ with domains $\Omega \restriction_U := \{(g,h) \in \Omega : g, h, g \cdot h \in U\}$ and $\Lambda \restriction_U := \{g \in \Lambda : g, g^{-1} \in U\}$, and with the group operations $\cdot, ()^{-1}$ being the restriction of the group operations of $G$ to $\Omega \restriction_U$, $\Lambda \restriction_U$ respectively. If $U$ is symmetric (in the sense that $g^{-1}$ is well-defined and lies in $U$ for all $g \in U$), then this restriction $G \restriction_U$ will also be symmetric. If $G$ is a global or local Lie group, then $G \restriction_U$ will also be a local Lie group. We will sometimes abuse notation and refer to the local group $G \restriction_U$ simply as $U$.

Thus, for instance, one can take the Euclidean space $\mathbf{R}^d$, and restrict it to a ball $B$ centred at the origin, to obtain an additive local group $\mathbf{R}^d \restriction_B$. In this group, two elements $x, y$ in $B$ have a well-defined sum $x + y$ only when

[1] Here we adopt the convention that any mathematical sentence involving an undefined operation is automatically false; thus, for instance, $g \cdot g^{-1} = 1$ is false unless $g \cdot g^{-1}$ is well-defined, so that $(g, g^{-1}) \in \Omega$.

their sum in $\mathbf{R}^d$ stays inside $B$. Intuitively, this local group behaves like the global group $\mathbf{R}^d$ as long as one is close enough to the identity element 0, but as one gets closer to the boundary of $B$, the group structure begins to break down.

It is natural to ask the question as to whether *every* local group arises as the restriction of a global group. The answer to this question is somewhat complicated, and can be summarised as "essentially yes in certain circumstances, but not in general"; see Chapter 15.

A key example of a local Lie group arises from pushing forward a Lie group via a coordinate chart near the origin:

**Example 2.1.3.** Let $G$ be a global or local Lie group of some dimension $d$, and let $\phi : U \to V$ be a smooth coordinate chart from a neighbourhood $U$ of the identity 1 in $G$ to a neighbourhood $V$ of the origin 0 in $\mathbf{R}^d$, such that $\phi$ maps 1 to 0. Then we can define a local group $\phi_* G \restriction_U$ which is the set $V$ (viewed as a smooth submanifold of $\mathbf{R}^d$) with the local group identity 0, the local group multiplication law $*$ defined by the formula

$$x * y := \phi(\phi^{-1}(x) \cdot \phi^{-1}(y))$$

defined whenever $\phi^{-1}(x), \phi^{-1}(y), \phi^{-1}(x) \cdot \phi^{-1}(y)$ are well-defined and lie in $U$, and the local group inversion law $()^{*-1}$ defined by the formula

$$x^{*-1} := \phi(\phi^{-1}(x)^{-1})$$

defined whenever $\phi^{-1}(x), \phi^{-1}(x)^{-1}$ are well-defined and lie in $U$. One easily verifies that $\phi_* G \restriction_U$ is a local Lie group. We will sometimes denote this local Lie group as $(V, *)$, to distinguish it from the additive local Lie group $(V, +)$ arising by restriction of $(\mathbf{R}^d, +)$ to $V$. The precise distinction between the two local Lie groups will in fact be a major focus of this section.

**Example 2.1.4.** Let $G$ be the Lie group $\mathrm{GL}_n(\mathbf{R})$, and let $U$ be the ball $U := \{g \in \mathrm{GL}_n(\mathbf{R}) : \|g - 1\|_{\mathrm{op}} < 1\}$. If we then let $V \subset M_n(\mathbf{R})$ be the ball $V := \{x \in M_n(\mathbf{R}) : \|x\|_{\mathrm{op}} < 1\}$ and $\phi$ be the map $\phi(g) := g - 1$, then $\phi$ is a smooth coordinate chart (after identifying $M_n(\mathbf{R})$ with $\mathbf{R}^{n \times n}$), and by the construction in the preceding exercise, $V = \phi_* G \restriction_U$ becomes a local Lie group with the operations

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

(defined whenever $x, y, x + y + xy$ all lie in $V$) and

$$x^{*-1} := (1 + x)^{-1} - 1 = x - x^2 + x^3 - \ldots$$

(defined whenever $x$ and $(1+x)^{-1}-1$ both lie in $V$). Note that this Lie group structure is not equal to the additive structure $(V, +)$ on $V$, nor is it equal to the multiplicative structure $(V, \cdot)$ on $V$ given by matrix multiplication,

which is one of the reasons why we use the symbol $*$ instead of $+$ or $\cdot$ for such structures.

Many (though not all) of the familiar constructions in group theory can be generalised to the local setting, though often with some slight additional subtleties. We will not systematically do so here, but we give a single such generalisation for now:

**Definition 2.1.5** (Homomorphism)**.** A *continuous homomorphism* $\phi : G \to H$ between two local groups $G, H$ is a continuous map from $G$ to $H$ with the following properties:

(i) $\phi$ maps the identity $1_G$ of $G$ to the identity $1_H$ of $H$: $\phi(1_G) = 1_H$.

(ii) If $g \in G$ is such that $g^{-1}$ is well-defined in $G$, then $\phi(g)^{-1}$ is well-defined in $H$ and is equal to $\phi(g^{-1})$.

(iii) If $g, h \in G$ are such that $g \cdot h$ is well-defined in $G$, then $\phi(g) \cdot \phi(h)$ is well-defined and equal to $\phi(g \cdot h)$.

A *smooth homomorphism* $\phi : G \to H$ between two local Lie groups $G, H$ is a continuous homomorphism that is also smooth.

A *(continuous) local homomorphism* $\phi : U \to H$ between two local groups $G, H$ is a continuous homomorphism from an open neighbourhood $U$ of the identity in $G$ to $H$. Two local homomorphisms are said to be *equivalent* if they agree on a (possibly smaller) open neighbourhood of the identity. One can of course define the notion of a smooth local homomorphism similarly.

It is easy to see that the composition of two continuous homomorphisms is again a continuous homomorphism, and that the identity map on a local group is automatically a continuous homomorphism; this gives the class of local groups the structure of a *category*. Similarly, the class of local Lie groups with their smooth homomorphisms is also a category.

**Example 2.1.6.** With the notation of Example 2.1.3, $\phi : U \to V$ is a smooth homomorphism from the local Lie group $G \restriction_U$ to the local Lie group $\phi_* G \restriction_U$. In fact, it is a smooth isomorphism, since $\phi^{-1} : V \to U$ provides the inverse homomorphism.

Let us say that a word $g_1 \dots g_n$ in a local group $G$ is *well-defined in* $G$ (or *well-defined*, for short) if every possible way of associating this word using parentheses is well-defined from applying the product operation. For instance, in order for $abcd$ to be well-defined, $((ab)c)d$, $(a(bc))d$, $(ab)(cd)$, $a(b(cd))$, and $a((bc)d)$ must all be well-defined. For instance, in the additive local group $\{-9, \dots, 9\}$ (with the group structure restricted from that of the integers $\mathbf{Z}$), $-2 + 6 + 5$ is not well-defined because one of the ways of

associating this sum, namely $-2+(6+5)$, is not well-defined (even though $(-2+6)+5$ is well-defined).

**Exercise 2.1.1** (Iterating the associative law)**.**

(i) Show that if a word $g_1 \dots g_n$ in a local group $G$ is well-defined, then all ways of associating this word give the same answer, and so we can uniquely evaluate $g_1 \dots g_n$ as an element in $G$.

(ii) Give an example of a word $g_1 \dots g_n$ in a local group $G$ which has two ways of being associated that are both well-defined, but give *different* answers. (*Hint:* The local associativity axiom prevents this from happening for $n \leq 3$, so try $n = 4$. A small discrete local group will already suffice to give a counterexample; verifying the local group axioms are easier if one makes the domain of definition of the group operations as small as one can get away with while still having the counterexample.)

**Exercise 2.1.2.** Show that the number of ways to associate a word $g_1 \dots g_n$ is given by the *Catalan number* $C_{n-1} := \frac{1}{n}\binom{2n-2}{n-1}$.

**Exercise 2.1.3.** Let $G$ be a local group, and let $m \geq 1$ be an integer. Show that there exists a symmetric open neighbourhood $U_m$ of the identity such that every word of length $m$ in $U_m$ is well-defined in $G$ (or more succinctly, $U_m^m$ is well-defined). (Note, though, that these words will usually only take values in $G$, rather than in $U_m$, and also the sets $U_m$ tend to become smaller as $m$ increases.)

## 2.2. Some differential geometry

To define the Lie algebra of a Lie group, we must first quickly recall some basic notions from differential geometry associated to smooth manifolds (which are not necessarily embedded in some larger Euclidean space, but instead exist intrinsically as abstract geometric structures). This requires a certain amount of abstract formalism in order to define things rigorously, though for the purposes of visualisation, it is more intuitive to view these concepts from a more informal geometric perspective.

We begin with the concept of the tangent space and related structures.

**Definition 2.2.1** (Tangent space)**.** Let $M$ be a smooth $d$-dimensional manifold. At every point $x$ of this manifold, we can define the *tangent space* $T_xM$ of $M$ at $x$. Formally, this tangent space can be defined as the space of all continuously differentiable curves $\gamma : I \to G$ defined on an open interval $I$ containing 0 with $\gamma(0) = x$, modulo the relation that two curves $\gamma_1, \gamma_2$ are

considered equivalent if they have the same derivative at 0, in the sense that

$$\frac{d}{dt}\phi(\gamma_1(t))|_{t=0} = \frac{d}{dt}\phi(\gamma_2(t))|_{t=0}$$

where $\phi : U \to V$ is a coordinate chart of $G$ defined in a neighbourhood of $x$; it is easy to see from the chain rule that this equivalence is independent of the actual choice of $\phi$. Using such a coordinate chart, one can identify the tangent space $T_xM$ with the Euclidean space $\mathbf{R}^d$, by identifying $\gamma$ with $\frac{d}{dt}\phi(\gamma(t))|_{t=0}$. One easily verifies that this gives $T_xM$ the structure of a $d$-dimensional vector space, in a manner which is independent of the choice of coordinate chart $\phi$. Elements of $T_xM$ are called *tangent vectors* of $M$ at $x$. If $\gamma : I \to G$ is a continuously differentiable curve with $\gamma(0) = x$, the equivalence class of $\gamma$ in $T_xM$ will be denoted $\gamma'(0)$.

The space $TM \coloneqq \bigcup_{x\in M}(\{x\} \times T_xM)$ of pairs $(x, v)$, where $x$ is a point in $M$ and $v$ is a tangent vector of $M$ at $x$, is called the *tangent bundle*.

If $\Phi : M \to N$ is a smooth map between two manifolds, we define the *derivative map* $D\Phi : TM \to TN$ to be the map defined by setting

$$D\Phi((x, \gamma'(0))) \coloneqq (\Phi(x), (\Phi \circ \gamma)'(0))$$

for all continously differentiable curves $\gamma : I \to G$ with $\gamma(0) = x$ for some $x \in M$; one can check that this map is well-defined. We also write $(\Phi(x), D\Phi(x)(v))$ for $D\Phi(x, v)$, so that for each $x \in M$, $D\Phi(x)$ is a map from $T_xM$ to $T_{\Phi(x)}N$. One can easily verify that this latter map is linear. We observe the *chain rule*[2]

$$D(\Psi \circ \Phi) = (D\Psi) \circ (D\Phi) \tag{2.1}$$

for any smooth maps $\Phi : M \to N$, $\Psi : N \to O$.

Observe that if $V$ is an open subset of $\mathbf{R}^d$, then $TV$ may be identified with $V \times \mathbf{R}^d$. In particular, every coordinate chart $\phi : U \to V$ of $M$ gives rise to a coordinate chart $D\phi : TU \to V \times \mathbf{R}^d$ of $TM$, which gives $TM$ the structure of a smooth $2d$-dimensional manifold.

**Remark 2.2.2.** Informally, one can think of a tangent vector $(x, v)$ as an infinitesimal vector from the point $x$ of $M$ to a nearby point $x + \varepsilon v + O(\varepsilon^2)$ on $M$, where $\varepsilon > 0$ is infinitesimally small; a smooth map $\phi$ then sends $x + \varepsilon v + O(\varepsilon^2)$ to $\phi(x) + \varepsilon D\phi(x)(v) + O(\varepsilon^2)$. One can make this informal perspective rigorous by means of *nonstandard analysis*, but we will not do so here.

Once one has the notion of a tangent bundle, one can define the notion of a smooth vector field:

[2]Indeed, one can view the tangent operator $T$ and the derivative operator $D$ together as a single *covariant functor* from the category of smooth manifolds to itself, although we will not need to use this perspective here.

**Definition 2.2.3** (Vector fields)**.** A *smooth vector field* on $M$ is a smooth map $X : M \to TM$ which is a right inverse for the projection map $\pi : TM \to M$, thus (by slight abuse of notation) $X$ maps $x$ to $(x, X(x))$ for some $X(x) \in T_x M$. The space of all smooth vector fields is denoted $\Gamma(TM)$. It is clearly a real vector space. In fact, it is a $C^\infty(M)$-module: given a smooth vector field $X \in \Gamma(TM)$ and a smooth function $f \in C^\infty(M)$ (i.e., a smooth map $f : M \to \mathbf{R}$), one can define the product $fX$ in the obvious manner: $fX(x) := f(x)X(x)$, and one easily verifies the axioms for a module.

Given a smooth function $f \in C^\infty(M)$ and a smooth vector field $X \in \Gamma(TM)$, we define the *directional derivative* $\nabla_X f \in C^\infty(M)$ of $f$ along $X$ by the formula

$$\nabla_X f(x) := \frac{d}{dt} f(\gamma(t))|_{t=0}$$

whenever $\gamma : I \to M$ is a continuously differentiable function with $\gamma(0) = x$ and $\gamma'(0) = X(x)$; one easily verifies that $\nabla_X f$ is well-defined and is an element of $C^\infty(M)$.

**Remark 2.2.4.** One can define $\nabla_X f$ in a more "coordinate free" manner as

$$\nabla_X f = \eta \circ Df \circ X,$$

where $\eta : T\mathbf{R} \to \mathbf{R}$ is the projection map to the second coordinate of $T\mathbf{R} \equiv \mathbf{R} \times \mathbf{R}$; one can also view $\nabla_X f$ as the *Lie derivative* of $f$ along $X$ (although, in most texts, the latter definition would be circular, because the Lie derivative is usually defined using the directional derivative).

**Remark 2.2.5.** If $V$ is an open subset of $\mathbf{R}^d$, a smooth vector field on $V$ can be identified with a smooth map $X : V \to \mathbf{R}^d$ from $V$ to $\mathbf{R}^d$. If $X : M \to TM$ is a smooth vector field on $M$ and $\phi : U \to V$ is a coordinate chart of $M$, then the *pushforward* $\phi_* X := D\phi \circ X \circ \phi^{-1} : V \to TV$ of $X$ by $\phi$ is a smooth vector field of $V$. Thus, in coordinates, one can view vector fields as maps from open subsets of $\mathbf{R}^d$ to $\mathbf{R}^d$. This perspective is convenient for quick and dirty calculations; for instance, in coordinates, the directional derivative $\nabla_X f$ is the same as the familiar directional derivative $X \cdot \nabla f$ from several variable calculus. If, however, one wishes to perform several changes of variable, then the more intrinsically geometric (and "coordinate-free") perspective outlined above can be more helpful.

There is a fundamental link between smooth vector fields and derivations of $C^\infty(M)$:

**Exercise 2.2.1** (Correspondence between smooth vector fields and derivations)**.** Let $M$ be a smooth manifold.

(i) If $X \in \Gamma(TM)$ is a smooth vector field, show that $\nabla_X : C^\infty(M) \to C^\infty(M)$ is a *derivation* on the (real) algebra $C^\infty(M)$, i.e., a (real) linear map that obeys the Leibniz rule

$$\nabla_X(fg) = f\nabla_X g + (\nabla_X f)g \tag{2.2}$$

for all $f, g \in C^\infty(M)$.

(ii) Conversely, if $d : C^\infty(M) \to C^\infty(M)$ is a derivation on $C^\infty(M)$, show that there exists a unique smooth vector field $X$ such that $d = \nabla_X$.

We see from the above exercise that smooth vector fields can be interpreted as a purely algebraic construction associated to the real algebra $C^\infty(M)$, namely as the space of derivations on that vector space. This can be useful for analysing the algebraic structure of such vector fields. Indeed, we have the following basic algebraic observation:

**Exercise 2.2.2** (Commutator of derivations is a derivation)**.** Let $d_1, d_2 : A \to A$ be two derivations on an algebra $A$. Show that the commutator $[d_1, d_2] := d_1 \circ d_2 - d_2 \circ d_1$ is also a derivation on $A$.

From the preceding two exercises, we can define the *Lie bracket* $[X, Y]$ of two vector fields $X, Y \in \Gamma(TM)$ by the formula

$$\nabla_{[X,Y]} := [\nabla_X, \nabla_Y].$$

This gives the space $\Gamma(TM)$ of smooth vector fields the structure of an (infinite-dimensional) *Lie algebra*:

**Definition 2.2.6** (Lie algebra)**.** A (real) Lie algebra is a real vector space $V$ (possibly infinite dimensional), together with a bilinear map $[,] : V \times V \to V$ which is anti-symmetric (thus $[X, Y] = -[Y, X]$ for all $X, Y \in V$, or equivalently $[X, X] = 0$ for all $X \in V$) and obeys the *Jacobi identity*

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \tag{2.3}$$

for all $X, Y, Z \in V$.

**Exercise 2.2.3.** If $M$ is a smooth manifold, show that $\Gamma(TM)$ (equipped with the Lie bracket) is a Lie algebra.

**Remark 2.2.7.** This is the abstract definition of a Lie algebra. A more concrete definition would be to let $V$ be a subspace of an algebra of operators, and to define the Lie bracket as the commutator. The relation between the two notions of a Lie algebra is explored in Chapter 13.

## 2.3. The Lie algebra of a Lie group

Let $G$ be a (global) Lie group. By definition, $G$ is then a smooth manifold, so we can thus define the tangent bundle $TG$ and smooth vector fields $X \in \Gamma(TG)$ as in the preceding section. In particular, we can define the tangent space $T_1 G$ of $G$ at the identity element 1.

If $g \in G$, then the left multiplication operation $\rho_g^{\text{left}} : x \mapsto gx$ is, by definition of a Lie group, a smooth map from $G$ to $G$. This creates a derivative map $D\rho_g^{\text{left}} : TG \to TG$ from the tangent bundle $TG$ to itself. We say that a vector field $X \in \Gamma(TG)$ is *left-invariant* if one has $(\rho_g^{\text{left}})_* X = X$ for all $g \in G$, or equivalently if $(D\rho_g^{\text{left}}) \circ X = X \circ \rho_g^{\text{left}}$ for all $g \in G$.

**Exercise 2.3.1.** Let $G$ be a (global) Lie group.

(i) Show that for every element $x$ of $T_1 G$ there is a unique left-invariant vector field $X \in \Gamma(TG)$ such that $X(1) = x$.

(ii) Show that the commutator $[X, Y]$ of two left-invariant vector fields is again a left-invariant vector field.

From the above exercise, we can identify the tangent space $T_1 G$ with the left-invariant vector fields on $TG$, and the Lie bracket structure on the latter then induces a Lie bracket (which we also call $[,]$) on $T_1 G$. The vector space $T_1 G$ together with this Lie bracket is then a (finite-dimensional) Lie algebra, which we call the *Lie algebra* of the Lie group $G$, and we write as $\mathfrak{g}$.

**Remark 2.3.1.** Informally, an element $x$ of the Lie algebra $\mathfrak{g}$ is associated with an infinitesimal perturbation $1 + \varepsilon x + O(\varepsilon^2)$ of the identity in the Lie group $G$. This intuition can be formalised fairly easily in the case of matrix Lie groups such as $\mathrm{GL}_n(\mathbf{C})$; for more abstract Lie groups, one can still formalise things using nonstandard analysis, but we will not do so here.

**Exercise 2.3.2.**

(i) Show that the Lie algebra $\mathfrak{gl}_n(\mathbf{C})$ of the general linear group $\mathrm{GL}_n(\mathbf{C})$ can be identified with the space $M_n(\mathbf{C})$ of $n \times n$ complex matrices, with the Lie bracket $[A, B] := AB - BA$.

(ii) Describe the Lie algebra $\mathfrak{u}_n(\mathbf{C})$ of the unitary group $\mathrm{U}_n(\mathbf{C})$.

(iii) Describe the Lie algebra $\mathfrak{su}_n(\mathbf{C})$ of the special unitary group $\mathrm{SU}_n(\mathbf{C})$.

(iv) Describe the Lie algebra $\mathfrak{o}_n(\mathbf{R})$ of the orthogonal $\mathrm{O}_n(\mathbf{R})$.

(v) Describe the Lie algebra $\mathfrak{so}_n(\mathbf{R})$ of the special orthogonal $\mathrm{SO}_n(\mathbf{R})$.

(vi) Describe the Lie algebra of the Heisenberg group $\begin{pmatrix} 1 & \mathbf{R} & \mathbf{R} \\ 0 & 1 & \mathbf{R} \\ 0 & 0 & 1 \end{pmatrix}$.

**Exercise 2.3.3.** Let $\phi : G \to H$ be a smooth homomorphism between (global) Lie groups. Show that the derivative map $D\phi(1_G)$ at the identity element $1_G$ is then a Lie algebra homomorphism from the Lie algebra $\mathfrak{g}$ of $G$ to the Lie algebra $\mathfrak{h}$ of $H$ (thus this map is linear and preserves the Lie bracket). (From this and the chain rule (2.1), we see that the map $\phi \mapsto D\phi(1_G)$ creates a covariant functor from the category of Lie groups to the category of Lie algebras.)

We have seen that every global Lie group gives rise to a Lie algebra. One can also associate Lie algebras to *local* Lie groups as follows:

**Exercise 2.3.4.** Let $G$ be a local Lie group. Let $U$ be a symmetric neighbourhood of the identity in $G$. (It is not difficult to see that at least one such neighbourhood exists.) Call a vector field $X \in \Gamma(TU)$ *left-invariant* if, for every $g \in U$, one has $(\rho_g^{\text{left}})_* X(g) = X(g)$, where $\rho_g^{\text{left}}$ is the left-multiplication map $x \mapsto gx$, defined on the open set $\{x \in U : gx \in U\}$ (where we adopt the convention that $gx \in U$ is shorthand for "$g \cdot x$ is well-defined and lies in $U$").

(i) Establish the analogue of Exercise 2.3.1 in this setting. Conclude that one can give $T_1 G$ the structure of a Lie algebra, which is independent of the choice of $U$.

(ii) Establish the analogue of Exercise 2.3.3 in this setting.

**Remark 2.3.2.** In the converse direction, it is also true that every finite-dimensional Lie algebra can be associated to either a local or a global Lie group; this is known as *Lie's third theorem*. However, this theorem is somewhat tricky to prove (particularly if one wants to associate the Lie algebra with a *global* Lie group), requiring the nontrivial algebraic tool of *Ado's theorem* (discussed in Section 13); see Exercise 2.5.6 below.

## 2.4. The exponential map

The *exponential map* $x \mapsto \exp(x)$ on the reals $\mathbf{R}$ (or its extension to the complex numbers $\mathbf{C}$) is of course fundamental to modern analysis. It can be defined in a variety of ways, such as the following:

(i) $\exp : \mathbf{R} \to \mathbf{R}$ is the differentiable map obeying the ODE $\frac{d}{dx}\exp(x) = \exp(x)$ and the initial condition $\exp(0) = 1$.

(ii) $\exp : \mathbf{R} \to \mathbf{R}$ is the differentiable map obeying the homomorphism property $\exp(x+y) = \exp(x)\exp(y)$ and the initial condition $\frac{d}{dx}\exp(x)|_{x=0} = 1$.

(iii) $\exp : \mathbf{R} \to \mathbf{R}$ is the limit of the functions $x \mapsto (1 + \frac{x}{n})^n$ as $n \to \infty$.

(iv) $\exp : \mathbf{R} \to \mathbf{R}$ is the limit of the infinite series $x \mapsto \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

We will need to generalise this map to arbitrary Lie algebras and Lie groups. In the case of matrix Lie groups (and matrix Lie algebras), one can use the matrix exponential, which can be defined efficiently by modifying definition (iv) above, and which was already discussed in Section 1. It is however difficult to use this definition for abstract Lie algebras and Lie groups. The definition based on (ii) will ultimately be the best one to use for the purposes of this text, but for foundational purposes (i) or (iii) is initially easier to work with. In most of the foundational literature on Lie groups and Lie algebras, one uses (i), in which case the existence and basic properties of the exponential map can be provided by the *Picard existence theorem* from the theory of ordinary differential equations. However, we will use (iii), because it relies less heavily on the smooth structure of the Lie group, and will therefore be more aligned with the spirit of Hilbert's fifth problem (which seeks to minimise the reliance of smoothness hypotheses whenever possible). Actually, for minor technical reasons it is slightly more convenient to work with the limit of $(1 + \frac{x}{2^n})^{2^n}$ rather than $(1 + \frac{x}{n})^n$.

We turn to the details. It will be convenient to work in local coordinates, and for applications to Hilbert's fifth problem it will be useful to "forget" almost all of the smooth structures. We make the following definition:

**Definition 2.4.1** ($C^{1,1}$ local group)**.** A $C^{1,1}$ *local group* is a local group $V$ that is an open neighbourhood of the origin 0 in a Euclidean space $\mathbf{R}^d$, with group identity 0, and whose group operation $*$ obeys the estimate

$$x * y = x + y + O(|x||y|) \tag{2.4}$$

for all sufficiently small $x, y$, where the implied constant in the $O()$ notation can depend on $V$ but is uniform in $x, y$.

**Example 2.4.2.** Let $G$ be a local Lie group of some dimension $d$, and let $\phi : U \to V$ be a smooth coordinate chart that maps a neighbourhood $U$ of the group identity 1 to a neighbourhood $V$ of the origin 0 in $\mathbf{R}^d$, with $\phi(1) = 0$. Then, as explained in Example 2.1.3, $V = (V, *) = \phi_* G \restriction_U$ is a local Lie group with identity 0; in particular, one has

$$0 * x = x * 0 = x.$$

From Taylor expansion (using the smoothness of $*$) we thus have (2.4) for sufficiently small $x, y$. Thus we see that every local Lie group generates a $C^{1,1}$ local group when viewed in coordinates.

**Remark 2.4.3.** In real analysis, a (locally) $C^{1,1}$ function is a function $f : U \to \mathbf{R}^m$ on a domain $U \subset \mathbf{R}^n$ which is continuously differentiable (i.e., in the regularity class $C^1$), and whose first derivatives $\nabla f$ are (locally) Lipschitz (i.e., in the regularity class $C^{0,1}$) the $C^{1,1}$ regularity class is

slightly weaker (i.e., larger) than the class $C^2$ of twice continuously differentiable functions, but much stronger than the class $C^1$ of singly continuously differentiable functions. See [**Ta2010**, §1.14] for more on these sorts of regularity classes. The reason for the terminology $C^{1,1}$ in the above definition is that $C^{1,1}$ regularity is essentially the minimal regularity for which one has the Taylor expansion

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + O(|x - x_0|^2)$$

for any $x_0$ in the domain of $f$, and any $x$ sufficiently close to $x_0$; note that the asymptotic (2.4) is of this form.

We now estimate various expressions in a $C^{1,1}$ local group.

**Exercise 2.4.1.** Let $V$ be a $C^{1,1}$ local group. Throughout this exercise, the implied constants in the $O()$ notation can depend on $V$, but not on parameters such as $x, y, \varepsilon, k, n$.

(i) Show that there exists an $\varepsilon > 0$ such that one has

$$x_1 * \cdots * x_k = x_1 + \cdots + x_k + O\left(\sum_{1 \leq i < j \leq k} |x_i||x_j|\right) \tag{2.5}$$

whenever $k \geq 1$ and $x_1, \ldots, x_k \in V$ are such that $\sum_{i=1}^k |x_i| \leq \varepsilon$, and the implied constant is uniform in $k$. Here and in the sequel we adopt the convention that a statement such as (2.5) is automatically false unless all expressions in that statement are well-defined. (*Hint:* Induct on $k$ using (2.4). It is best to replace the asymptotic $O()$ notation by explicit constants $C$ in order to ensure that such constants remain uniform in $k$.) In particular, one has the crude estimate

$$x_1 * \cdots * x_k = O(\sum_{i=1}^k |x_i|)$$

under the same hypotheses as above.

(ii) Show that one has

$$x^{*-1} = -x + O(|x|^2)$$

for $x$ sufficiently close to the origin.

(iii) Show that

$$x * y * x^{*-1} * y^{*-1} = O(|x||y|)$$

for $x, y$ sufficiently close to the origin. (*Hint:* First show that $x * y = y * x + O(|x||y|)$, then express $x * y$ as the product of $x * y * x^{*-1} * y^{*-1}$ and $y * x$.)

(iv) Show that
$$x * y * x^{*-1} = y + O(|x||y|)$$
whenever $x, y$ are sufficiently close to the origin.

(v) Show that
$$y * x^{*-1}, x^{*-1} * y = O(|x - y|)$$
whenever $x, y$ are sufficiently close to the origin.

(vi) Show that there exists an $\varepsilon > 0$ such that
$$x_1 * \cdots * x_k = y_1 * \cdots * y_k + O\left(\sum_{i=1}^{k} |x_i - y_i|\right)$$
whenever $k \geq 1$ and $x_1, \ldots, x_k, y_1, \ldots, y_k$ are such that
$$\sum_{i=1}^{k} |x_i|, \sum_{j=1}^{k} |y_i| \leq \varepsilon.$$

(vii) Show that there exists an $\varepsilon > 0$ such that
$$\frac{1}{2}|n||x - y| \leq |x^{*n} - y^{*n}| \leq 2|n||x - y|$$
for all $n \in \mathbf{Z}$ and $x, y \in \mathbf{R}^d$ such that $|nx|, |ny| \leq \varepsilon$, where $x^{*n} = x * \cdots * x$ is the product of $n$ copies of $x$ (assuming of course that this product is well-defined) for $n \geq 0$, and $x^{*-n} := (x^{*n})^{*-1}$.

(viii) Show that there exists an $\varepsilon > 0$ such that
$$(xy)^{*n} = x^{*n} y^{*n} + O\left(|n|^2 |x||y|\right)$$
for all $n \in \mathbf{Z}$ and $x, y \in \mathbf{R}^d$ such that $|nx|, |ny| \leq \varepsilon$. (*Hint:* Do the case when $n$ is positive first. In that case, express $x^{*-n} * (xy)^{*n}$ as the product of $n$ conjugates of $y$ by various powers of $x$.)

We can now define the *exponential map* $\exp : V' \to V$ on this $C^{1,1}$ local group by defining
$$\exp(x) := \lim_{n \to \infty} \left(\frac{1}{2^n} x\right)^{*2^n} \tag{2.6}$$
for any $x$ in a sufficiently small neighbourhood $V'$ of the origin in $V$.

**Exercise 2.4.2.** Let $V$ be a local $C^{1,1}$ group.

(i) Show that if $V'$ is a sufficiently small neighbourhood of the origin in $V$, then the limit in (2.6) exists for all $x \in V'$. (*Hint:* Use the previous exercise to estimate the distance between $(\frac{1}{2^n} x)^{*2^n}$ and $(\frac{1}{2^{n+1}} x)^{*2^{n+1}}$.) Establish the additional estimate
$$\exp(x) = x + O(|x|^2). \tag{2.7}$$

(ii) Show that if $\gamma : I \to G$ is a smooth curve with $\gamma(0) = 1$, and $\gamma'(0)$ is sufficiently small, then

$$\exp(\gamma'(0)) = \lim_{n\to\infty} \gamma(1/2^n)^{*2^n}.$$

(iii) Show that for all sufficiently small $x, y$, one has the bilipschitz property

$$|(\exp(x) - \exp(y)) - (x - y)| \leq \frac{1}{2}|x - y|.$$

Conclude, in particular, that for $V'$ sufficiently small, exp is a homeomorphism between $V'$ and an open neighbourhood $\exp(V')$ of the origin. (*Hint:* To show that $\exp(V')$ contains a neighbourhood of the origin, use (2.7) and the contraction mapping theorem.)

(iv) Show that

$$\exp(sx) * \exp(tx) = \exp((s + t)x) \tag{2.8}$$

for $s, t \in \mathbf{R}$ and $x \in \mathbf{R}^d$ with $sx, tx$ sufficiently small. (*Hint:* First handle the case when $s, t \in \mathbf{Z}[\frac{1}{2}]$ are dyadic numbers.)

(v) Show that for any sufficiently small $x, y \in \mathbf{R}^d$, one has

$$\exp(x + y) = \lim_{n\to\infty} (\exp(x/2^n) * \exp(y/2^n))^{*2^n}. \tag{2.9}$$

Then conclude the stronger estimate

$$\exp(x + y) = \lim_{n\to\infty} (\exp(x/n) * \exp(y/n))^{*n}. \tag{2.10}$$

(vi) Show that for any sufficiently small $x, y \in \mathbf{R}^d$, one has

$$\exp(x + y) = \exp(x) * \exp(y) + O(|x||y|).$$

(*Hint:* Use the previous part, as well as Exercise 2.4.1(viii).)

Let us say that a $C^{1,1}$ local group is *radially homogeneous* if one has

$$sx * tx = (s + t)x \tag{2.11}$$

whenever $s, t \in \mathbf{R}$ and $x \in \mathbf{R}^d$ are such that $sx, tx$ are sufficiently small. (In particular, this implies that $x^{*-1} = -x$ for sufficiently small $x$.) From the above exercise, we see that any $C^{1,1}$ local group $V$ can be made into a radially homogeneous $C^{1,1}$ local group $V'$ by first restricting to an open neighbourhood $\exp(V')$ of the identity, and then applying the logarithmic homeomorphism $\exp^{-1}$. Thus:

**Corollary 2.4.4.** *Every $C^{1,1}$ local group has a neighbourhood of the identity which is isomorphic (as a topological group) to a radially homogeneous $C^{1,1}$ local group.*

Now we study the exponential map on global Lie groups. If $G$ is a global Lie group, and $\mathfrak{g}$ is its Lie algebra, we define the exponential map $\exp : \mathfrak{g} \to G$ on a global Lie group $G$ by setting

$$\exp(\gamma'(0)) := \lim_{n\to\infty} \gamma(1/2^n)^{2^n}$$

whenever $\gamma : I \to G$ is a smooth curve with $\gamma(0) = 1$.

**Exercise 2.4.3.** Let $G$ be a global Lie group.

(i) Show that the exponential map is well-defined. (*Hint:* First handle the case when $\gamma'(0)$ is small, using the previous exercise, then bootstrap to larger values of $\gamma'(0)$.)

(ii) Show that for all $x, y \in \mathfrak{g}$ and $s, t \in \mathbf{R}$, one has

$$\exp(sx)\exp(tx) = \exp((s+t)x) \tag{2.12}$$

and

$$\exp(x+y) = \lim_{n\to\infty} (\exp(x/n)\exp(y/n))^n. \tag{2.13}$$

(*Hint:* Again, begin with the case when $x, y$ are small.)

(iii) Show that the exponential map is continuous.

(iv) Show that for each $x \in \mathfrak{g}$, the function $t \mapsto \exp(tx)$ is the unique homomorphism from $\mathbf{R}$ to $G$ that is differentiable at $t = 0$ with derivative equal to $x$.

**Proposition 2.4.5** (Lie's first theorem)**.** *Let $G$ be a Lie group. Then the exponential map is smooth. Furthermore, there is an open neighbourhood $U$ of the origin in $\mathfrak{g}$ and an open neighbourhood $V$ of the identity in $G$ such that the exponential map* exp *is a diffeomorphism from $U$ to $V$.*

**Proof.** We begin with the smoothness. From the homomorphism property we see that

$$\frac{d}{dt}\exp(tx) = \left(\rho^{\text{left}}_{\exp(tx)}\right)_* x$$

for all $x \in \mathfrak{g}$ and $t \in \mathbf{R}$. If $x$ and $t$ are sufficiently small, and one uses a coordinate chart $\phi$ near the origin, the function $f(t,x) := \phi(\exp(tx))$ then satisfies an ODE of the form

$$\frac{d}{dt} f(t,x) = F(f(t,x), x)$$

for some smooth function $F$, with initial condition $f(0,x) = 0$; thus by the fundamental theorem of calculus we have

$$f(t,x) = \int_0^t F(f(t',x),x)\, dt'. \tag{2.14}$$

Now let $k \geq 0$. An application of the contraction mapping theorem (in the function space $L_t^\infty C_x^k$ localised to small region of spacetime) then shows that $f$ lies in $L_t^\infty C_x^k$ for small enough $t, x$, and by further iteration of the integral equation we then conclude that $f(t,x)$ is $k$ times continuously differentiable for small enough $t, x$. By (2.8) we then conclude that exp is smooth everywhere.

Since

$$\frac{d}{dt} \exp(tx)|_{t=0} = x$$

we see that the derivative of the exponential map at the origin is the identity map on $\mathfrak{g}$. The second claim of the proposition thus follows from the inverse function theorem. □

In view of this proposition, we see that given a vector space basis $X_1, \ldots, X_d$ for the Lie algebra $\mathfrak{g}$, we may obtain a smooth coordinate chart $\phi : U \to V$ for some neighbourhood $U$ of the identity and neighbourhood $V$ of the origin in $\mathbf{R}^d$ by defining

$$\tilde{\phi}(\exp(t_1 X_1 + \cdots + t_d X_d)) := (t_1, \ldots, t_d)$$

for sufficiently small $t_1, \ldots, t_d \in \mathbf{R}$. These are known as *exponential coordinates of the first kind*. Although we will not use them much here, we also note that there are *exponential coordinates of the second kind*, in which the expression $\exp(t_1 X_1 + \cdots + t_d X_d)$ is replaced by the slight variant $\exp(t_1 X_1) \ldots \exp(t_d X_d)$.

Using exponential coordinates of the first kind, we see that we may identify a local piece $U$ of the Lie group $G$ with the radially homogeneous $C^{1,1}$ local group $V$. In the next section, we will analyse such radially homogeneous $C^{1,1}$ groups further. For now, let us record some easy consequences of the existence of exponential coordinates. Define a *one-parameter subgroup* of a topological group $G$ to be a continuous homomorphism $\phi : \mathbf{R} \to G$ from $\mathbf{R}$ to $G$.

**Exercise 2.4.4** (Classification of one-parameter subgroups)**.** Let $G$ be a Lie group. For any $X \in \mathfrak{g}$, show that the map $t \mapsto \exp(tX)$ is a one-parameter subgroup. Conversely, if $\phi : \mathbf{R} \to G$ is a one-parameter subgroup, there exists a unique $X \in \mathfrak{g}$ such that $\phi(t) = \exp(tX)$ for all $t \in \mathbf{R}$. (*Hint:* Mimic the proof of Proposition 1.0.1.)

**Proposition 2.4.6** (Weak regularity implies strong regularity)**.** *Let $G, H$ be global Lie groups, and let $\Phi : G \to H$ be a continuous homomorphism. Then $\Phi$ is smooth.*

**Proof.** Since $\Phi$ is a continuous homomorphism, it maps one-parameter subgroups of $G$ to one-parameter subgroups of $H$. Thus, for every $X \in \mathfrak{g}$, there

exists a unique element $L(X) \in \mathfrak{h}$ such that

$$\Phi(\exp(tX)) = \exp(tL(X))$$

for all $t \in \mathbf{R}$. In particular, we see that $L$ is homogeneous: $L(sX) = sL(X)$ for all $X \in \mathfrak{g}$ and $s \in \mathbf{R}$. Next, we observe using (2.9) and the fact that $\Phi$ is a continuous homomorphism that for any $X, Y \in \mathfrak{g}$ and $t \in \mathbf{R}$, one has

$$\begin{aligned}\Phi(\exp(t(X+Y))) &= \Phi\left(\lim_{n\to\infty} (\exp(tX/2^n)\exp(tY/2^n))^{2^n}\right)\\ &= \lim_{n\to\infty} (\Phi(\exp(tX/2^n))\Phi(\exp(tY/2^n)))^{2^n}\\ &= \lim_{n\to\infty} (\exp(tL(X)/2^n)\exp(tL(Y)/2^n))^{2^n}\\ &= \exp(t(L(X)+L(Y)))\end{aligned}$$

and thus $L$ is additive:

$$L(X+Y) = L(X) + L(Y).$$

We conclude that $L$ is a linear transformation from the finite-dimensional vector space $\mathfrak{g}$ to the finite-dimensional vector space $\mathfrak{h}$. In particular, $L$ is smooth. On the other hand, we have

$$\Phi(\exp(X)) = \exp(L(X)).$$

Since $\exp : \mathfrak{g} \to G$ and $\exp : \mathfrak{h} \to H$ are diffeomorphisms near the origin, we conclude that $\Phi$ is smooth in a neighbourhood of the identity. Using the homomorphism property (and the fact that the group operations are smooth for both $G$ and $H$) we conclude that $\Phi$ is smooth everywhere, as required. □

This fact has a pleasant corollary:

**Corollary 2.4.7** (Uniqueness of Lie structure)**.** *Any (global) topological group can be made into a Lie group in at most one manner. More precisely, given a topological group $G$, there is at most one smooth structure one can place on $G$ that makes the group operations smooth.*

**Proof.** Suppose for the sake of contradiction that one could find two different smooth structures on $G$ that make the group operations smooth, leading to two different Lie groups $G', G''$ based on $G$. The identity map from $G'$ to $G''$ is a continuous homomorphism, and hence smooth by the preceding proposition; similarly for the inverse map from $G''$ to $G'$. This implies that the smooth structures coincide, and the claim follows. □

Note that a general high-dimensional topological manifold may have more than one smooth structure, which may even be nondiffeomorphic to each other (as the example of *exotic spheres* [**Mi1956**] demonstrates), so this corollary is not entirely vacuous.

**Exercise 2.4.5.** Let $G$ be a connected (global) Lie group, let $H$ be another (global) Lie group, and let $\Phi : G \to H$ be a continuous homomorphism (which is thus smooth by Proposition 2.4.6). Show that $\Phi$ is uniquely determined by the derivative map $D\Phi(1) : \mathfrak{g} \to \mathfrak{h}$. In other words, if $\Phi' : G \to H$ is another continuous homomorphism with $D\Phi(1) = D\Phi'(1)$, then $\Phi = \Phi'$. (*Hint:* First prove this in a small neighbourhood of the origin. What group does this neighbourhood generate?) What happens if $G$ is not connected?

**Exercise 2.4.6** (Weak regularity implies strong regularity, local version)**.** Let $G, H$ be local Lie groups, and let $\Phi : G \to H$ be a continuous homomorphism. Show that $\Phi$ is smooth in a neighbourhood of the identity in $G$.

Now we can establish the final stage, at least, of the program outlined in Figure 1:

**Exercise 2.4.7** (Local Lie implies Lie)**.** Let $G$ be a global topological group. Suppose that there is an open neighbourhood $U$ of the identity such that the local group $G \restriction_U$ can be given the structure of a local Lie group. Show that $G$ can be given the structure of a global Lie group. (*Hint:* We already have at least one coordinate chart on $G$; translate it around to create an atlas of such charts. To show compatibility of the charts and global smoothness of the group, one needs to show that the conjugation maps $x \mapsto gxg^{-1}$ are smooth near the origin for any $g \in G$. To prove this, use Exercise 2.4.6.)

## 2.5. The Baker-Campbell-Hausdorff formula

We now study radially homogeneous $C^{1,1}$ local groups in more detail, in particular, filling in some of the last few steps in the program in Figure 1. We will show

**Theorem 2.5.1** (Baker-Campbell-Hausdorff formula, qualitative version)**.** *Let $V \subset \mathbf{R}^d$ be a radially homogeneous $C^{1,1}$ local group. Then the group operation $*$ is real analytic near the origin. In particular, after restricting $V$ to a sufficiently small neighbourhood of the origin, one obtains a local Lie group.*

We will in fact give a more precise formula for $*$, known as the *Baker-Campbell-Haudorff-Dynkin formula*, in the course of proving Theorem 2.5.1. This formula is usually proven just for Lie groups, but it turns out that the proof of the formula extends without much difficulty to the $C^{1,1}$ local group setting (the main difference being that continuous operations, such as Riemann integrals, have to be replaced by discrete counterparts, such as Riemann sums).

**Remark 2.5.2.** In the case where $V$ comes from viewing a general linear group $\mathrm{GL}_n(\mathbf{C})$ in local exponential coordinates, the group operation $*$ is given by $x * y = \log(\exp(x)\exp(y))$ for sufficiently small $x, y \in M_n(\mathbf{C})$. Thus, a corollary of Theorem 2.5.1 is that this map is real analytic.

We begin the proof of Theorem 2.5.1. Throughout this section, $V \subset \mathbf{R}^d$ is a fixed radially homogeneous $C^{1,1}$ local group. We will need some variants of the basic bound (2.4).

**Exercise 2.5.1** (Lipschitz bounds)**.** If $x, y, z \in V$ are sufficiently small, establish the bounds

$$x * y = x + y + O(|x+y||y|), \tag{2.15}$$

$$x * y = x + y + O(|x+y||x|), \tag{2.16}$$

$$x * y = x * z + O(|y-z|), \tag{2.17}$$

and

$$y * x = z * x + O(|y-z|). \tag{2.18}$$

(*Hint:* To prove (2.15), start with the identity $(x * y) * (-y) = x$.)

Now we exploit the radial homogeneity to describe the conjugation operation $y \mapsto x * y * (-x)$ as a linear map:

**Lemma 2.5.3** (Adjoint representation)**.** *For all $x$ sufficiently close to the origin, there exists a linear transformation $\mathrm{Ad}_x : \mathbf{R}^d \to \mathbf{R}^d$ such that $x * y * (-x) = \mathrm{Ad}_x(y)$ for all $y$ sufficiently close to the origin.*

**Remark 2.5.4.** Using the matrix example from Remark 2.5.2, we are asserting here that

$$\exp(x)\exp(y)\exp(-x) = \exp(\mathrm{Ad}_x(y))$$

for some linear transform $\mathrm{Ad}_x(y)$ of $y$, and all sufficiently small $x, y$. Indeed, using the basic matrix identity $\exp(AxA^{-1}) = A\exp(x)A^{-1}$ for invertible $A$ (coming from the fact that the conjugation map $x \mapsto AxA^{-1}$ is a continuous ring homomorphism) we see that we may take $\mathrm{Ad}(x) = \exp(x)y\exp(-x)$ here.

**Proof.** Fix $x$. The map $y \mapsto x * y * (-x)$ is continuous near the origin, so it will suffice to establish additivity, in the sense that

$$x * (y+z) * (-x) = (x * y * (-x)) + (x * z * (-x))$$

for $y, z$ sufficiently close to the origin.

Let $n$ be a large natural number. Then from (2.11) we have

$$(y+z) = \left(\frac{1}{n}y + \frac{1}{n}z\right)^{*n}.$$

Conjugating this by $x$, we see that

$$\begin{aligned} x*(y+z)*(-x) &= \left(x*\left(\frac{1}{n}y+\frac{1}{n}z\right)*(-x)\right)^n \\ &= n\left(x*\left(\frac{1}{n}y+\frac{1}{n}z\right)*(-x)\right). \end{aligned}$$

But from (2.4) we have

$$\frac{1}{n}y + \frac{1}{n}z = \frac{1}{n}y * \frac{1}{n}z + O\left(\frac{1}{n^2}\right)$$

and thus (by Exercise 2.5.1)

$$x*\left(\frac{1}{n}y+\frac{1}{n}z\right)*(-x) = x*\frac{1}{n}y*\frac{1}{n}z*(-x) + O\left(\frac{1}{n^2}\right).$$

But if we split $x*\frac{1}{n}y*\frac{1}{n}z*(-x)$ as the product of $x*\frac{1}{n}y*(-x)$ and $x*\frac{1}{n}z*(-x)$ and use (2.4), we have

$$x*\frac{1}{n}y*\frac{1}{n}z*(-x) = x*\frac{1}{n}y*(-x) + x*\frac{1}{n}z*(-x) + O\left(\frac{1}{n^2}\right).$$

Putting all this together we see that

$$\begin{aligned} x*(y+z)*(-x) &= n\left(x*\frac{1}{n}y*(-x) + x*\frac{1}{n}z*(-x) + O\left(\frac{1}{n^2}\right)\right) \\ &= x*y*(-x) + x*z*(-x) + O\left(\frac{1}{n}\right); \end{aligned}$$

sending $n \to \infty$ we obtain the claim. □

From (2.4) we see that

$$\|\operatorname{Ad}_x - I\|_{\mathrm{op}} = O(|x|)$$

for $x$ sufficiently small. Also from the associativity property we see that

$$\operatorname{Ad}_{x*y} = \operatorname{Ad}_x \operatorname{Ad}_y \tag{2.19}$$

for all $x, y$ sufficiently small. Combining these two properties (and using (2.15)) we conclude in particular that

$$\|\operatorname{Ad}_x - \operatorname{Ad}_y\|_{\mathrm{op}} = O(|x-y|) \tag{2.20}$$

for $x, y$ sufficiently small. Thus we see that Ad is a (locally) continuous linear representation. In particular, $t \mapsto \operatorname{Ad}_{tx}$ is a (locally) continuous

homomorphism into a linear group, and so (by Proposition 1.0.1) we have the *Hadamard lemma*

$$\mathrm{Ad}_x = \exp(\mathrm{ad}_x)$$

for all sufficiently small $x$, where $\mathrm{ad}_x : \mathbf{R}^d \to \mathbf{R}^d$ is the linear transformation

$$\mathrm{ad}_x = \frac{d}{dt} \mathrm{Ad}_{tx} \,|_{t=0}.$$

From (2.19), (2.20), (2.4) we see that

$$\mathrm{Ad}_{tx} \mathrm{Ad}_{ty} = \mathrm{Ad}_{t(x+y)} + O(|t|^2)$$

for $x, y, t$ sufficiently small, and so by the product rule we have

$$\mathrm{ad}_{x+y} = \mathrm{ad}_x + \mathrm{ad}_y .$$

Also, we clearly have $\mathrm{ad}_{tx} = t\,\mathrm{ad}_x$ for $x, t$ small. Thus we see that $\mathrm{ad}_x$ is linear in $x$, and so we have

$$\mathrm{ad}_x\, y = [x, y] \tag{2.21}$$

for some bilinear form $[,] : \mathbf{R}^d \to \mathbf{R}^d$.

One can show that this bilinear form in fact defines a Lie bracket (i.e., it is anti-symmetric and obeys the Jacobi identity), but for now, all we need is that it is manifestly real analytic (since all bilinear forms are polynomial and thus analytic). In particular, $\mathrm{ad}_x$ and $\mathrm{Ad}_x$ depend analytically on $x$.

We now give an important approximation to $x * y$ in the case when $y$ is small:

**Lemma 2.5.5.** *For $x, y$ sufficiently small, we have*

$$x * y = x + F(\mathrm{Ad}_x)y + O(|y|^2)$$

*where*

$$F(z) := \frac{z \log z}{z - 1}.$$

**Proof.** If we write $z := x * y - x$, then $z = O(|y|)$ (by (2.4)) and

$$(-x) * (x + z) = y.$$

We will shortly establish the approximation

$$(-x) * (x + z) = \frac{1 - \exp(-\mathrm{ad}_x)}{\mathrm{ad}_x} z + O(|z|^2); \tag{2.22}$$

inverting

$$\frac{1 - \exp(-\mathrm{ad}_x)}{\mathrm{ad}_x} = \frac{\mathrm{Ad}_x - 1}{\mathrm{Ad}_x \log \mathrm{Ad}_x}$$

we obtain the claim.

It remains to verify (2.22). Let $n$ be a large natural number. We can expand the left-hand side of (2.22) as a telescoping series

$$\sum_{j=0}^{n-1} \left(-\frac{j+1}{n}x\right) * \left(\frac{j+1}{n}x + \frac{j+1}{n}z\right) - \left(-\frac{j}{n}x\right) * \left(\frac{j}{n}x + \frac{j}{n}z\right). \tag{2.23}$$

Using (2.11), the first summand can be expanded as

$$\left(-\frac{j}{n}x\right) * \left(-\frac{x}{n}\right) * \left(\frac{x}{n} + \frac{z}{n}\right) * \left(\frac{j}{n}x + \frac{j}{n}z\right).$$

From (2.15) one has $\left(-\frac{x}{n}\right) * \left(\frac{x}{n} + \frac{z}{n}\right) = \frac{z}{n} + O(\frac{|z|}{n^2})$, so by (2.17), (2.18) we can write the preceding expression as

$$\left(-\frac{j}{n}x\right) * \frac{z}{n} * \left(\frac{j}{n}x + \frac{j}{n}z\right) + O\left(\frac{|z|}{n^2}\right)$$

which by definition of Ad can be rewritten as

$$\left(\mathrm{Ad}_{-\frac{j}{n}x}\frac{z}{n}\right) * \left(-\frac{j}{n}x\right) * \left(\frac{j}{n}x + \frac{j}{n}z\right) + O\left(\frac{|z|}{n^2}\right). \tag{2.24}$$

From (2.15) one has

$$\left(-\frac{j}{n}x\right) * \left(\frac{j}{n}x + \frac{j}{n}z\right) = O(|z|)$$

while from (2.20) one has $\mathrm{Ad}_{-\frac{j}{n}x}\frac{z}{n} = O(|z|/n)$, hence from (2.4) we can rewrite (2.24) as

$$\mathrm{Ad}_{-\frac{j}{n}x}\frac{z}{n} + \left(-\frac{j}{n}x\right) * \left(\frac{j}{n}x + \frac{j}{n}z\right) + O\left(\frac{|z|^2}{n}\right) + O\left(\frac{|z|}{n^2}\right).$$

Inserting this back into (2.23), we can thus write the left-hand side of (2.22) as

$$\left(\sum_{j=0}^{n-1} \mathrm{Ad}_{-\frac{j}{n}x}\frac{z}{n}\right) + O(|z|^2) + O\left(\frac{|z|}{n}\right).$$

Writing $\mathrm{Ad}_{-\frac{j}{n}x} = \exp\left(-\frac{j}{n}\,\mathrm{ad}_x\right)$, and then letting $n \to \infty$, we conclude (from the convergence of the Riemann sum to the Riemann integral) that

$$(-x) * (x+z) = \int_0^1 \exp(-t\,\mathrm{ad}_x)z\; dt + O(|z|^2)$$

and the claim follows. □

**Remark 2.5.6.** In the matrix case, the key computation is to show that

$$\exp(-x)\exp(x+z) = 1 + \frac{1-\exp(-\mathrm{ad}_x)}{\mathrm{ad}_x}z + O(|z|^2).$$

To see this, we can use the fundamental theorem of calculus to write the left-hand side as

$$1 + \int_0^1 \frac{d}{dt}(\exp(-tx)\exp(t(x+z)))\ dt.$$

Since $\frac{d}{dt}\exp(-tx) = \exp(-tx)(-x)$ and $\frac{d}{dt}\exp(t(x+z)) = (x+z)\exp(t(x+z))$, we can rewrite this as

$$1 + \int_0^1 \exp(-tx)z\exp(t(x+z))\ dt.$$

Since $\exp(t(x+z)) = \exp(tx) + O(|z|)$, this becomes

$$1 + \int_0^1 \exp(-tx)z\exp(tx)\ dt + O(|z|^2);$$

since $\exp(-tx)z\exp(tx) = \exp(-t\,\mathrm{ad}_x)z$, we obtain the desired claim.

We can integrate the above formula to obtain an exact formula for $*$:

**Corollary 2.5.7** (Baker-Campbell-Hausdorff-Dynkin formula)**.** *For $x, y$ sufficiently small, one has*

$$x * y = x + \int_0^1 F(\mathrm{Ad}_x \mathrm{Ad}_{ty})y\ dt.$$

The right-hand side is clearly real analytic in $x$ and $y$, and Theorem 2.5.1 follows.

**Proof.** Let $n$ be a large natural number. We can express $x * y$ as the telescoping sum

$$x + \sum_{j=0}^{n-1} x * \left(\frac{j+1}{n}y\right) - x * \left(\frac{j}{n}y\right).$$

From (2.11) followed by Lemma 2.5.5 and (2.19), one has

$$\begin{aligned} x * \left(\frac{j+1}{n}y\right) &= x * \left(\frac{j}{n}y\right) * \frac{y}{n} \\ &= x * \left(\frac{j}{n}y\right) + F\left(\mathrm{Ad}_x \mathrm{Ad}_{\frac{j}{n}y}\right)\frac{y}{n} + O\left(\frac{1}{n^2}\right). \end{aligned}$$

We conclude that

$$x * y = x + \frac{1}{n}\sum_{j=0}^{n-1} F\left(\mathrm{Ad}_x \mathrm{Ad}_{\frac{j}{n}y}\right)y + O\left(\frac{1}{n}\right).$$

Sending $n \to \infty$, so that the Riemann sum converges to a Riemann integral, we obtain the claim. □

**Remark 2.5.8.** It is not immediately obvious from this formula alone why $*$ should be associative. A derivation of associativity from the Baker-Campbell-Hausdorff-Dynkin formula is given in Chapter 14.

**Exercise 2.5.2.** Use the Taylor-type expansion

$$F(z) = 1 - \frac{1/z-1}{2} + \frac{(1/z-1)^2}{3} - \frac{(1/z-1)^3}{4} + \dots$$

to obtain the explicit expansion

$$x * y = x + \sum_{n=0}^{\infty} \frac{(-1)^m}{n+1} \sum_{\substack{r_i,s_i \geq 0 \\ (r_i,s_i) \neq (0,0)}} \frac{(\mathrm{ad}_y)^{r_1}(\mathrm{ad}_x)^{s_1} \dots (\mathrm{ad}_y)^{r_n}(\mathrm{ad}_x)^{s_n}}{r_1! s_1! \dots r_n! s_n! (r_1 + \dots + r_n + 1)} y$$

where $m := n + r_1 + \dots + r_n + s_1 + \dots + s_n + 1$, and show that the series is absolutely convergent for $x, y$ small enough. Invert this to obtain the alternate expansion

$$x * y = y + \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} \sum_{\substack{r_i,s_i \geq 0 \\ (r_i,s_i) \neq (0,0)}} \frac{(\mathrm{ad}_x)^{r_1}(\mathrm{ad}_y)^{s_1} \dots (\mathrm{ad}_x)^{r_n}(\mathrm{ad}_y)^{s_n}}{r_1! s_1! \dots r_n! s_n! (r_1 + \dots + r_n + 1)} x.$$

**Exercise 2.5.3.** Let $V$ be a radially homogeneous $C^{1,1}$ local group. By Theorem 2.5.1, an open neighbourhood of the origin in $V$ has the structure of a local Lie group, and thus by Exercise 2.3.4 is associated to a Lie algebra. Show that this Lie algebra is isomorphic to $\mathbf{R}^d$ and the Lie bracket $[,]$ is given by (2.19). Note that this establishes *a posteriori* the fact that the bracket $[,]$ occurring in (2.19) is anti-symmetric and obeys the Jacobi identity.

We now record some consequences of the Baker-Campbell-Hausdorff formula.

**Exercise 2.5.4** (Lie groups are analytic)**.** Let $G$ be a global Lie group. Show that $G$ is a real analytic manifold (i.e., one can find an atlas of smooth coordinate charts whose transition maps are all real analytic), and that the group operations are also real analytic (i.e., they are real analytic when viewed in the above-mentioned coordinate charts). Furthermore, show that any continuous homomorphism between Lie groups is also real analytic.

**Exercise 2.5.5** (Lie's second theorem)**.** Let $G, H$ be global Lie groups, and let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. Show that there exists an

open neighbourhood $U$ of the identity in $G$ and a homomorphism $\Phi : U \to H$ from the local Lie group $G \lfloor_U$ to $H$ such that $D\Phi(1) = \phi$. If $G$ is connected and simply connected, show that one can take $U$ to be all of $G$.

**Exercise 2.5.6** (Lie's third theorem)**.** *Ado's theorem* asserts that every finite-dimensional Lie algebra is isomorphic to a subalgebra of $\mathfrak{g}l_n(\mathbf{R})$ for some $n$. This (somewhat difficult) theorem and its proof is discussed in Chapter 13. Assuming Ado's theorem as a "black box", conclude the following claims:

(i) (Lie's third theorem, local version) Every finite-dimensional Lie algebra is isomorphic to the Lie algebra of some local Lie group.

(ii) Every local or global Lie group has a neighbourhood of the identity that is isomorphic to a local *linear* Lie group (i.e., a local Lie group contained in $\mathrm{GL}_n(\mathbf{R})$ or $\mathrm{GL}_n(\mathbf{C})$ for some $n$).

(iii) (Lie's third theorem, global version) Every finite-dimensional Lie algebra $\mathfrak{g}$ is isomorphic to the Lie algebra of some global Lie group. (*Hint:* From (i) and (ii), one may identify $\mathfrak{g}$ with the Lie algebra of a local linear Lie group. Now consider the space of all smooth curves in the ambient linear group that are everywhere "tangent" to this local linear Lie group modulo "homotopy", and use this to build the global Lie group.)

(iv) (Lie's third theorem, simply connected version) Every finite-dimensional Lie algebra $\mathfrak{g}$ is isomorphic to the Lie algebra of some global connected, simply connected Lie group. Furthermore, this Lie group is unique up to isomorphism.

(v) Show that every local Lie group $G$ has a neighbourhood of the identity that is isomorphic to a neighbourhood of the identity of a global connected, simply connected Lie group. Furthermore, this Lie group is unique up to isomorphism.

**Remark 2.5.9.** One does not need the full strength of Ado's theorem to establish conclusion (i) of the above exercise. Indeed, it suffices to show that the operation $*$ defined in Exercise 2.5.2 is associative near the origin. To do this, it suffices to verify associativity in the sense of formal power series; and then by abstract nonsense one can lift up to the free Lie algebra on $d$ generators, and then down to the free *nilpotent* Lie algebra on $d$ generators and of some arbitrary finite step $s$, which one can verify to be a finite-dimensional Lie algebra. Applying Ado's theorem for the special case of nilpotent Lie algebras (which is easier to establish than the general case of Ado's theorem, as discussed in Chapter 13), one can identify this nilpotent Lie algebra with a subalgebra of $\mathfrak{g}_n(\mathbf{R})$ for some $n$, and then one can argue

as in the above exercise to conclude. See also Chapter 14 for an alternate way to establish associativity of $*$. However, I do not know how to establish conclusions (ii), (iii) or (iv) without using Ado's theorem in full generality (and (ii) is in fact *equivalent* to this theorem, at least in characteristic 0).

**Remark 2.5.10.** Lie's three theorems can be interpreted as establishing an *equivalence* between three different categories: the category of finite-dimensional Lie algebras; the category of local Lie groups (or more precisely, the category of local Lie group *germs*, formed by identifying local Lie groups that are identical near the origin); and the category of global connected, simply connected Lie groups. See Chapter 15 for further discussion.

The fact that we were able to establish the Baker-Campbell-Hausdorff formula at the $C^{1,1}$ regularity level will be useful for the purposes of proving results related to Hilbert's fifth problem. In particular, we have the following criterion for a group to be Lie (very much in accordance with the rigidity principle from the introduction):

**Lemma 2.5.11** (Criterion for Lie structure)**.** *Let $G$ be a topological group. Then $G$ is Lie if and only if there is a neighbourhood of the identity in $G$ which is isomorphic (as a topological group) to a $C^{1,1}$ local group.*

This gives the last three steps of the program in Figure 1.

**Proof.** The "only if" direction is trivial. For the "if" direction, combine Corollary 2.4.4 with Theorem 2.5.1 and Exercise 2.4.7. □

**Remark 2.5.12.** Informally, Lemma 2.5.11 asserts that $C^{1,1}$ regularity can automatically be upgraded to smooth ($C^\infty$) or even real analytic ($C^\omega$) regularity for topological groups. In contrast, note that a locally Euclidean group has neighbourhoods of the identity that are isomorphic to a "$C^0$ local group" (which is the same concept as a $C^{1,1}$ local group, but without the asymptotic (2.4)). Thus we have reduced Hilbert's fifth problem to the task of boosting $C^0$ regularity to $C^{1,1}$ regularity, rather than that of boosting $C^0$ regularity to $C^\infty$ regularity.

**Exercise 2.5.7.** Let $G$ be a Lie group with Lie algebra $\mathfrak{g}$. For any $X, Y \in \mathfrak{g}$, show that

$$\exp([X,Y]) = \lim_{n\to\infty} (\exp(X/n)\exp(Y/n)\exp(-X/n)\exp(-Y/n))^{n^2}.$$