VERY BASIC LIE THEORY

ROGER HOWE
Department of Mathematics, Yale University, New Haven, CT 06520

Lie theory, the theory of Lie groups, Lie algebras and their applications, is a fundamental part of mathematics. Since World War II it has been the focus of a burgeoning research effort, and is now seen to touch a tremendous spectrum of mathematical areas, including classical, differential, and algebraic geometry, topology, ordinary and partial differential equations, complex analysis (one and several variables), group and ring theory, number theory, and physics, from classical to quantum and relativistic.

It is impossible in a short space to convey the full compass of the subject, but we will cite some examples. An early major success of Lie theory, occurring when the subject was still in its infancy, was to provide a systematic understanding of the relationship between Euclidean geometry and the newer geometries (hyperbolic non-Euclidean or Lobachevskian, Riemann's elliptic geometry, and projective geometry) that had arisen in the 19th century. This led Felix Klein to enunciate his Erlanger Programm [**Kl**] for the systematic understanding of geometry. The principle of Klein's program was that geometry should be understood as the study of quantities left invariant by the action of a group on a space. Another development in which Klein was involved was the Uniformization Theorem [**Be**] for Riemann surfaces. This theorem may be understood as saying that every connected two-manifold is a double coset space of the isometry group of one of the 3 (Euclidean, hyperbolic, elliptic) standard 2-dimensional geometries. (See also the recent article [**F**] in this MONTHLY.) Three-manifolds are much more complex than two manifolds, but the intriguing work of Thurston [**Th**] has gone a long way toward showing that much of their structure can be understood in a way analogous to the 2-dimensional situation in terms of coset spaces of certain Lie groups.

More or less contemporary with the final proof of the Uniformization Theorem was Einstein's [**E**] invention of the special theory of relativity and its instatement of the Lorentz transformation as a basic feature of the kinematics of space-time. Einstein's intuitive treatment of relativity was followed shortly by a more sophisticated treatment by Minkowski [**Mk**] in which Lorentz transformations were shown to constitute a certain Lie group, the isometry group of an indefinite Riemannian metric on $\mathbb{R}^4$. Similarly, shortly after Heisenberg [**Hg**] introduced his famous Commutation Relations in quantum mechanics, which underlie his Uncertainty Principle, Hermann Weyl [**W**] showed they could be interpreted as the structure relations for the Lie algebra of a certain two-step nilpotent Lie group. As the group-theoretical underpinnings of physics became better appreciated, some physicists, perhaps most markedly Wigner [**Wg**], in essence advocated extending Klein's Erlanger Programm to physics. Today, indeed, symmetry principles based on Lie theory are a standard tool and a major source of progress in theoretical physics. Quark theory [**Dy**], in particular, is primarily a (Lie) group-theoretical construct.

These examples could be multiplied many times. The applications of Lie theory are astonishing in their pervasiveness and sometimes in their unexpectedness. The articles of Borel [**Bo2**] and Dyson [**Dy**] mention some. The recent article of Proctor [**Pr**] in this MONTHLY discusses an application to combinatorics. Some points of contact of Lie theory with the undergraduate curriculum are listed in §7.

The article of Proctor also illustrates the need to broaden understanding of Lie theory. Proctor did not feel he could assume knowledge of basic Lie theoretic facts. Though hardly an unknown subject, Lie theory is poorly known in comparison to its importance. Especially since it provides

---

Roger Howe received his Ph.D. from the University of California at Berkeley in 1969. His advisor was Calvin C. Moore. He was at the State University of New York at Stony Brook from 1969 to 1974, and since then has been at Yale University. His main research interests are in group representation theory and harmonic analysis, both pure and applied.

unity of methods and viewpoints in the many subjects to which it relates, its wide dissemination seems worthwhile. Yet it has barely penetrated the undergraduate curriculum, and it is far from universally taught in graduate programs.

Part of the reason for the pedagogy gap is that standard treatments [**A**], [**Ch**], [**He**] of the foundations of Lie theory involve substantial prerequisites, including the basic theory of differentiable manifolds, some additional differential geometry, and the theory of covering spaces. This approach tends to put a course in Lie theory, when available, in the second year of graduate study, after specialization has already begun. While a complete discussion of Lie theory does require fairly elaborate preparation, a large portion of its essence is accessible on a much simpler level, appropriate to advanced undergraduate instruction. This paper attempts to present the theory at that level. It presupposes only a knowledge of point set topology and calculus in normed vector spaces. In fact, for the Lie theory proper, only normed vector spaces are necessary. This simplification is achieved by not considering general or abstract Lie groups, but only groups concretely realized as groups of matrices. Since such groups provide the great bulk of significant examples of Lie groups, for many purposes this restriction is unimportant.

The essential phenomenon of Lie theory, to be explicated in the rest of this paper, is that one may associate in a natural way to a Lie group G its Lie algebra $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is first of all a vector space and secondly is endowed with a bilinear nonassociative product called the *Lie bracket* or *commutator* and usually denoted $[\,,]$. Amazingly, the group G is almost completely determined by $\mathfrak{g}$ and its Lie bracket. Thus for many purposes one can replace G with $\mathfrak{g}$. Since G is a complicated nonlinear object and $\mathfrak{g}$ is just a vector space, it is usually vastly simpler to work with $\mathfrak{g}$. Otherwise intractable computations may become straightforward linear algebra. This is one source of the power of Lie theory.

The basic object mediating between Lie groups and Lie algebras is the one-parameter group. Just as an abstract group is a coherent system of cyclic groups, a Lie group is a (very) coherent system of one-parameter groups. The purpose of the first two sections, therefore, is to provide some general philosophy about one-parameter groups. Section 1 provides background on homeomorphism groups, and one-parameter groups are defined in a general context in §2. Discussion of Lie groups proper begins in §3. Technically it is independent of §§1 and 2; but these sections will, I hope, give some motivation for reading on. Those who need no motivation or dislike philosophy may go directly to §3. There one-parameter groups of linear transformations are defined and are described by means of the exponential map on matrices. In §4 the exponential map is studied, and the commutator bracket makes its appearance. Section 5 is the heart of the paper. It defines and gives examples of matrix groups, the class of Lie groups considered in this paper. Then it defines Lie algebras, and shows that every matrix group can be associated to a Lie algebra which is related to its group in a close and precise way. The main statement is Theorem 17, and Theorem 19 and Corollary 20 are important complements. Finally §6 ties up some loose ends and §7, as noted, describes some connections of Lie theory with the standard curriculum.

Bibliographical note: The arguments of sections 3, 4 and 5 are very close to those given by von Neumann [**Nn**] in his 1929 paper on Hilbert's 5th problem. A modern development of basic Lie theory which incorporates these results is [**Go**].

1. Homeomorphism Groups

In this section, we use the standard terminology of general topology, as for example in [**Ke**].

Let X be a set. Then the collection $\mathrm{Bi}(X)$ of bijections from X to itself is a group with composition of mappings as the group law. Now suppose X is in fact a topological space. Then the set $\mathrm{Hm}(X)$ of homeomorphisms from X to itself is a subgroup of $\mathrm{Bi}(X)$. It seems natural to try to topologize $\mathrm{Hm}(X)$. The topology should of course reflect how $\mathrm{Hm}(X)$ acts on X, so that maps close to the identity move points very little. But a topology on $\mathrm{Hm}(X)$ should also be consistent with the group structure of $\mathrm{Hm}(X)$. More precisely and generally, given a group G, if it is to be made into a topological space in a manner consistent with its group structure, the topology it is

given should satisfy two conditions.

(1.1) (i) The multiplication map $(g_1, g_2) \to g_1 g_2$ from $G \times G$ to G should be continuous.
(ii) The inverse map $g \to g^{-1}$ from G to G should be continuous.

A topology on G satisfying these two compatibility criteria is called a *group topology*. A group endowed with a group topology is called a *topological group*. Some standard treatments of topological groups are **[Hn]** and **[P]**.

In short, then, we would like to make $\mathrm{Hm}(G)$ into a topological group. This is not so satisfactorily done for completely general X, but if X is locally compact Hausdorff, there is a nice topology on $\mathrm{Hm}(X)$, known as the *compact-open topology*. Before defining it, we make some general observations about group topologies. These will simplify the definition.

Given a group G and an element $g \in G$, define λ_g, *left-translation by* g, and ρ_g, *right-translation by* g, to be the maps

$$\lambda_g : G \to G \qquad \rho_g : G \to G \tag{1.2}$$

given by

$$\lambda_g(g') = gg' \qquad \rho_g(g') = g'g^{-1}.$$

For $U \subseteq G$, set

$$gU = \lambda_g(U) \qquad Ug = \rho_{g^{-1}}(U). \tag{1.3}$$

LEMMA 1. *Let G be a topological group, and $g \in G$.*

(a) *The map $\lambda_g : G \to G$ is a homeomorphism. Similarly $\rho_g : G \to G$ is a homeomorphism.*

(b) *If $U \subseteq G$ is a neighborhood of the identity 1_G of G, then gU and Ug are neighborhoods of g. Similarly if $V \subseteq G$ is a neighborhood of g, then $g^{-1}V$ and Vg^{-1} are neighborhoods of 1_G.*

Proof. One checks from the definition of λ_g that λ is a homeomorphism, i.e.,

$$\lambda_g \circ \lambda_h = \lambda_{gh} \tag{1.4}$$

for $g, h \in G$. It follows directly from the condition (1.1) (i) that λ_g is continuous. Likewise, the map $\lambda_{g^{-1}}$ is also continuous. From (1.4) one concludes that

$$\lambda_{g^{-1}} = (\lambda_g)^{-1}. \tag{1.5}$$

Hence λ_g is continuous with continuous inverse, that is, a homeomorphism. The proof for ρ_g is essentially identical.

Since $\lambda_g(1_G) = g$ and $\lambda_g(U) = gU$ by definition, part b) follows since the homeomorphic image of an open set is open. ■

COROLLARY 2. *A group topology is determined by its system of neighborhoods of the identity.*

Proof. Indeed, a topology on G is determined by the collection of neighborhood systems of each point of G. But according to part (b) of the lemma, for a group topology, the system of neighborhoods around a point $g \in G$ is determined by the system of neighborhoods around 1_G. ■

Let us call a topology on G such that all λ_g and ρ_g are homeomorphisms a *homogeneous topology*. Lemma 1 says group topologies are homogeneous. Evidently Corollary 2 applies to all homogeneous topologies, not only group topologies. Thus an obvious question is what conditions must a neighborhood system at the identity satisfy in order that the associated homogeneous topology be a group topology? This question has a simple answer.

LEMMA 3. *A homogeneous topology on a group G is a group topology if and only if the system of neighborhoods of 1_G satisfies conditions* (a) *and* (b) *below.*

(a) *If U is a neighborhood of 1_G, there is another neighborhood V of 1_G such that $V \subseteq U^{-1}$, where*

$$U^{-1} = \{g^{-1} : g \in U\}. \tag{1.6}$$

(b) *If U is a neighborhood of 1_G, there are other neighborhoods V, W of 1_G such that $VW \subseteq U$, where*

$$VW = \{gh : g \in V, h \in W\}. \tag{1.7}$$

Proof. The conditions (a) and (b) are clearly necessary for a topology to be a group topology, since they are one way of stating that the inverse and multiplication maps are continuous at 1_G. We will check this for condition (b). In order for multiplication to be continuous at $1_G \times 1_G$, given a neighborhood $U \subseteq G$ of 1_G, we must find a neighborhood $U' \subseteq G \times G$ of $1_G \times 1_G$ such that for any point (g, g') of U', the product gg' is in U. But by definition of the product topology on $G \times G$, any neighborhood U' of $1_G \times 1_G$ contains a product $V \times W$, where $V, W \subseteq G$ are both neighborhoods of 1_G. But the image of $V \times W$ under the multiplication map is just the set VW defined in (1.7). So condition (b) amounts to continuity of multiplication at the point $1_G \times 1_G \in G \times G$.

Thus to complete the lemma we need to show that if the multiplication and inverse maps are continuous at the identity, and if the topology on G is homogeneous, then they are continuous everywhere. Let U be a neighborhood of 1_G. Then since λ_g is a homeomorphism, gU is a neighborhood of g, and we may write

$$(gu)^{-1} = u^{-1}g^{-1} = \rho_g(u^{-1}) = \rho_g\left(\left(\lambda_{g^{-1}}(gu)\right)^{-1}\right) \quad u \in U. \tag{1.8}$$

Thus on gU, the inverse map is a composition of $\lambda_{g^{-1}}$, the inverse map on U, and ρ_g. Since $\lambda_{g^{-1}}$ and ρ_g are continuous, and $\lambda_{g^{-1}}$ takes g to 1_G, and the inverse map is continuous at 1_G, we see that the inverse map is continuous at g also. The proof that multiplication is continuous everywhere is analogous and is left as an exercise. ■

We return to the question of topologizing $\mathrm{Hm}(X)$. Corollary 2 allows us to save work in our definition of the topology on $\mathrm{Hm}(X)$ by only defining neighborhoods of the identity map 1_X on X, and declaring by fiat all left or right translates of these neighborhoods also to be open sets. Lemma 3 tells us what we must check to know our definition yields a group topology.

From now on, we take X to be a locally compact Hausdorff space. Let $C \subseteq X$ be compact, and let $O \supseteq C$ be open. Define

$$U(C, O) = \{h \in \mathrm{Hm}(X) : h(C) \subseteq O, h^{-1}(C) \subseteq O\}. \tag{1.9}$$

If $\{C_i\}$, $1 \leqslant i \leqslant n$, are compact subsets of X, and $\{O_i\}$ are open subsets of X such that $C_i \subseteq O_i$, set

$$U(\{C_i\}\{O_i\}) = \bigcap_{i=1}^{n} U(C_i, O_i). \tag{1.10}$$

DEFINITION. Let X be a locally compact Hausdorff space. The *compact-open* topology on $\mathrm{Hm}(X)$ is the homogeneous topology such that a base for the neighborhoods of 1_X consists of the sets $U(\{C_i\}, \{O_i\})$ of equation (1.10).

PROPOSITION 4. *The compact-open topology on* $\mathrm{Hm}(X)$ *is a Hausdorff group topology.*

Proof. Since we have decreed the compact open topology to be homogeneous, we need only check the conditions of Lemma 3 to show it is a group topology. Condition (a) is automatic since the sets $U(C, O)$ are defined to be invariant under the inverse map on $\mathrm{Hm}(X)$. Let us check condition (b). If U_i, V_i and W_i are neighborhoods of 1_X such that $V_iW_i \subseteq U_i$, then evidently

$$\left(\bigcap_i V_i\right)\left(\bigcap_i W_i\right) \subseteq \bigcap_i U_i.$$

Hence since the sets (1.10) are intersections of the sets $U(C, O)$ of (1.9), it will be enough to check condition (b) with the neighborhood U of the form $U = U(C, O)$. Since X is locally compact Hausdorff, we can by a standard separation theorem (cf. [**Ke**, Chap. 5, Theorem 18]) find an open

$O' \subseteq X$ such that the closure C' of O' is compact, and

$$C \subseteq O' \subseteq C' \subseteq O.$$

Then set $V = W = U(C', O) \cap U(C, O')$. If $h_1, h_2 \in V$, we find

$$h_1 \circ h_2(C) = h_1(h_2(C)) \subseteq h_1(O') \subseteq h_1(C') \subseteq O$$

and similarly for $(h_1 \circ h_2)^{-1} = h_2^{-1} \circ h_1^{-1}$. Thus $h_1 \circ h_2 \in U(C, O) = U$, or $VW \subseteq U$ as was to be shown, and the compact-open topology is a group topology on $\mathrm{Hm}(X)$.

To show that a group topology is Hausdorff is a fairly simple matter. We record the relevant observation as a separate result.

LEMMA 5. *Let G be a topological group. Let $H \subseteq G$ be the intersection of all neighborhoods of 1_G. Then H is a normal subgroup of G. Further, G is Hausdorff if and only if $H = \{1_G\}$.*

Proof. Suppose $h_1, h_2 \in H$. Given a neighborhood U of 1_G, we can find neighborhoods V, W of 1_G such that $VW \subseteq U$. Since $h_1 \in V$ and $h_2 \in W$, we see that $h_1 h_2 \in U$. Hence $h_1 h_2 \in H$ also. In similar fashion, one sees that $h_1^{-1} \in H$. Hence H is a group. Since the conjugate gUg^{-1} of a neighborhood of 1_G is again a neighborhood of 1_G, we see that H is also normal in G.

Suppose $H = 1_G$. Then given $g \in G$, we can find a neighborhood U of 1_G such that $g \notin U$. Let V, W be neighborhoods of 1_G such that $VW \subseteq U$. Then gV^{-1} and W are neighborhoods of g and of 1_G, respectively, and are disjoint. Now consider any two points $g_1, g_2 \in G$. Set $g = g_1^{-1} g_2$, and apply the argument above. Translating on the left by g_1, we find $g_1 W$ and $g_2 V^{-1}$ are disjoint neighborhoods of g_1 and g_2, respectively. Hence G is Hausdorff. ■

From Lemma 5 we see Proposition 4 will be proved if we produce for each $h \neq 1_X$ in $\mathrm{Hm}(X)$ a compact C and open O such that $h \notin U(C, O)$. Choose $x \in X$ such that $h(x) \neq x$. Then evidently $h \notin U(\{x\}, X - \{h(x)\})$. ■

REMARK. In fact the compact-open topology on $\mathrm{Hm}(X)$ is better than Proposition 4 indicates. It is complete with respect to an appropriate uniform structure ([**Ke**, Chap. 6]). Also, if X is second countable (hence metrizable), then $\mathrm{Hm}(X)$ is also second countable and metrizable.

2. One-Parameter Groups: Flows and Differential Equations

The real number system $\mathbb{R}$ equipped with addition and its familiar topology is, as the reader may easily check, a topological group.

DEFINITION. A *one-parameter group* of homeomorphisms of (the locally compact Hausdorff space) X is a continuous homomorphism

$$\varphi : \mathbb{R} \to \mathrm{Hm}(X). \tag{2.1}$$

It will be convenient to denote the image under φ of t by φ_t rather than $\varphi(t)$. Thus $\{\varphi_t\}$ is a family of homeomorphisms of X satisfying the rule

$$\varphi_t \circ \varphi_s = \varphi_{t+s} \qquad t, s \in \mathbb{R}. \tag{2.2}$$

Since for each t the map φ_t acts on X, a one-parameter group of homeomorphisms of X is also called an $\mathbb{R}$-action on X, or an action by $\mathbb{R}$ on X.

Given a one-parameter group φ_t of homeomorphisms of X, we can define a map

$$\begin{aligned} &\Phi : \mathbb{R} \times X \to X, \\ &\Phi(t, x) = \varphi_t(x). \end{aligned} \tag{2.3}$$

The fact that $t \to \varphi_t$ is a homomorphism is captured by the identities

$$\begin{aligned} &\text{(i)}\ \Phi(0, x) = x, \\ &\text{(ii)}\ \Phi(s, \Phi(t, x)) = \Phi(s + t, x). \end{aligned} \tag{2.4}$$

The continuity of φ is reflected in the continuity of Φ. We state this fact formally. It will perhaps also shed some light on the significance of the compact-open topology on $\mathrm{Hm}(X)$.

LEMMA 6. *Let $\Phi : \mathbb{R} \times X \to X$ be a map. For $t \in \mathbb{R}$, define $\varphi_t : X \to X$ by formula* (2.3) (ii). *Then $\{\varphi_t\}$ is a one-parameter group of homeomorphisms if and only if*

(a) *Φ satisfies identities* (2.4) *and*

(b) *Φ is continuous.*

REMARK. According to this lemma, if our goal were simply to define a one parameter group of homeomorphisms in the quickest way, we could short-circuit the whole discussion of §1 and simply define a one-parameter group of homeomorphisms as a map Φ satisfying the conditions of the lemma. However, that approach seemed unduly formalistic.

Proof. It is a straightforward computation to verify that the identities (2.4) guarantee that for each t the map φ_t is in $\mathrm{Bi}(X)$ and $t \to \varphi_t$ is a homomorphism. Also it is obvious that the maps φ_t will be in $\mathrm{Hm}(X)$ if and only if Φ is continuous in x for each fixed t. Thus the main thrust of the lemma is that $t \to \varphi_t$ is continuous from $\mathbb{R}$ to $\mathrm{Hm}(X)$ if and only if Φ is jointly continuous in t and x. Let us verify this.

Suppose Φ is continuous. Let $C \subseteq X$ be compact, and $O \subseteq X$ be open, with $C \subseteq O$. Choose any $x \in C$. By identity (2.4)(i), the point $(0, x) \in \mathbb{R} \times X$ is in $\Phi^{-1}(0)$. By continuity of Φ a neighborhood of $(0, x)$ is contained in $\Phi^{-1}(0)$. This means there is a neighborhood N of x in X, and $\delta > 0$, depending on x, N, and O, such that $\Phi(t, y) \in 0$ for $y \in N$ and $|t| < \delta$. In other words $\varphi_t(y) \in O$ for $|t| < \delta$ and $y \in N$. Since C is compact, we can find a finite number of $x_i \in C$ such that the associated neighborhoods N_i cover C. Suppose then that $\Phi(t, y_i) \in O$ for $y_i \in N_i$ and $|t| < \delta_i$. Set $\delta = \min \delta_i$. Then we have $\Phi(t, c) \in O$ for all $c \in C$ and $|t| < \delta$. In other words, $\varphi_t \in U(C, O)$ for $|t| < \delta$. Clearly, by repeating this argument for any finite collection of compact C_i's and open O_i's containing them, we can show that $\varphi_t \in U(\{C_i\}, \{O_i\})$ for all sufficiently small t. This shows that $t \to \varphi_t$ is continuous at the origin in $\mathbb{R}$. But now we appeal to the following lemma.

LEMMA 7. *Let $\varphi : G \to H$ be a homomorphism between topological groups. Then φ is continuous if and only if φ is continuous at 1_G.*

The proof of this lemma is left as an exercise to the reader, who will recognize in it the same spirit that informs Lemmas 2, 3, and 5.

To finish Lemma 6, we must show that the continuity of $t \to \varphi_t$ implies continuity of Φ. Choose $(t, x) \in \mathbb{R} \times X$, and set $y = \Phi(t, x)$. Let V be a neighborhood of y. Since φ_t is continuous, we can find a neighborhood W of x, with compact closure $\overline{W}$, such that $\varphi_t(\overline{W}) \subseteq V$. Since φ_t is continuous in t, we can find $\varepsilon > 0$ so that $\varphi_s \in U(\varphi_t(\overline{W}), V)$ for $|s| < \varepsilon$. But then if $(t', w) \in (t - \varepsilon, t + \varepsilon) \times W$, we have

$$\Phi(t', w) = \Phi(t' - t, \Phi(t, w)) = \varphi_{t'-t}(\Phi(t, w)) \in \varphi_{t'-t}(\varphi_t(W)) \subseteq V.$$

In other words $(t - \varepsilon, t + \varepsilon) \times W \subseteq \Phi^{-1}(V)$. Since V was an arbitrary neighborhood of y, we see Φ is continuous at (t, x). Since (t, x) is arbitrary, we see Φ is continuous. ∎

Consider a one-parameter group φ_t of homeomorphisms of X and the associated map Φ defined by formula (2.3). The map Φ is a function of two variables, t and x, and the maps φ_t are obtained from Φ by temporarily fixing t and letting x vary. If on the other hand we fix x and let t vary, we get a map $t \to \Phi(t, x) = \varphi_t(x)$ which defines a continuous curve in X, traced by the moving point $\varphi_t(x)$. Thus as t varies, each point of x moves continuously inside X, and various points move in a coherent fashion, so that we can form a mental picture of them flowing through X, each point along its individual path. For this reason, a one-parameter group of homeomorphisms of X is also sometimes called a *flow* on X.

The notion of a flow is closely related to the theory of differential equations. Indeed, let $X = \mathbb{R}^n$, and write

$$x = (x_1, \ldots, x_n) \qquad x \in \mathbb{R}^n, x_i \in \mathbb{R}.$$

Then

$$\Phi(t, x) = \Phi(t, x_1, x_2, \ldots, x_n) = (\Phi_1(t, x_1, \ldots, x_n), \ldots, \Phi_n(t, x_1, \ldots, x_n))$$

is a function from $\mathbb{R}^{n+1}$ to $\mathbb{R}^n$. Suppose Φ is not merely continuous, but differentiable. Define

$$f: \mathbb{R}^n \to \mathbb{R}^n$$

by

$$f(x) = \frac{\partial \Phi}{\partial t}(t, x)\bigg|_{t=0}. \tag{2.5}$$

If we differentiate (2.4) (ii) with respect to s, and set $s = 0$, we obtain

$$\frac{\partial \Phi(t, x)}{\partial t} = f(\Phi(t, x)). \tag{2.6}$$

In other words, for fixed x, the map $y_x(t) = \Phi(t, x)$ is a solution of the system of differential equations

$$\frac{dy}{dt} = f(y), \tag{2.7}$$

or

$$\frac{dy_i}{dt} = f_i(y_1, \ldots, y_n), \quad \text{for } 1 \leqslant i \leqslant n.$$

The solution y_x of (2.7) is the solution of (2.7) with initial condition $y_x(0) = x$.

The system (2.7) may be pictured geometrically as follows. At each point $y \in \mathbb{R}^n$, one draws the vector $f(y) = (f_1(y), f_2(y), \ldots, f_n(y))$. This gives a family of vectors which vary smoothly as y varies; such a family is called a *vector field*. A solution of the system (2.7) is a parametrized curve $c(t)$ in $\mathbb{R}^n$, such that at each point $c(t)$ of the curve the tangent vector $c'(t)$ is the pre-assigned vector $f(c(t))$. The 2-dimensional system

$$\frac{d(x, y)}{dt} = (-y, x)$$

whose solutions are the circles

$$(x(t), y(t)) = (a\cos(\theta_0 + t), a\sin(\theta_0 + t))$$

is illustrated in Fig. 1.

Suppose on the other hand that for each x we have a solution $y_x(t)$ of the system (2.7) with initial condition $y_x(0) = x$. For $s \in R$, consider the function

$$y_{x,s}(t) = y_x(t + s).$$

Differentiation of $y_{x,s}$ shows it also is a solution of the system (2.7), evidently with initial value $y_{x,s}(0) = y_x(s)$. The uniqueness part of the Existence and Uniqueness Theorem for ordinary differential equations [**L**], [**HS**], [**R**], therefore implies that

$$y_x(s + t) = y_{x,s}(t) = y_{y_x(s)}(t). \tag{2.8}$$

If we then set

$$\Phi(t, x) = y_x(t),$$

we find that identity (2.8) translates into identity (2.4) (ii). Of course, the initial condition $y_x(0) = x$ is just identity (2.4) (i). It follows that $\varphi_t(x) = y_x(t)$ defines a one-parameter group of

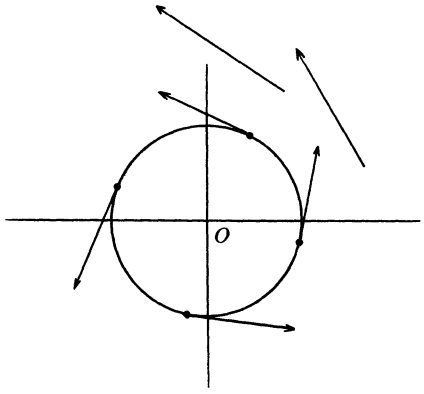

FIG. 1.

homeomorphisms of $\mathbb{R}^n$. For the system of Fig. 1, the map φ_t is just rotation through an angle of t radians.

In summary, we have seen that a (smooth) one-parameter group of diffeomorphisms of $\mathbb{R}^n$ yields solutions of a system of differential equations of the form (2.7), and, conversely, a solution (for all time and all x) of system (2.7) yields a one-parameter group. The two constructs, solutions of systems of ordinary differential equations, and one-parameter groups, thus provide two different points of view on the same mathematical phenomenon. In other words, the notion of one-parameter group provides a geometric and global way of looking at the solutions of a system of ordinary differential equations. As such, it suggests ways of attacking and obtaining information about ordinary differential equations, and it provides a link between systems of ordinary differential equations and more complex geometric objects such as the Lie groups and Lie algebras discussed in the following sections.

3. One-Parameter Groups of Linear Transformations

In this section, we show how one-parameter groups of linear transformations of a vector space can be described using the exponential map on matrices.

Let V be a finite dimensional real vector space. Let $\mathrm{End}(V)$ denote the algebra of linear maps from V to itself, and let $\mathrm{GL}(V)$ denote the group of invertible linear maps from V to itself. The usual name for $\mathrm{GL}(V)$ is the *general linear group* of V. If $V = \mathbb{R}^n$, then $\mathrm{End}(V) = M_n(\mathbb{R})$, the $n \times n$ matrices, and $\mathrm{GL}(V) = \mathrm{GL}_n(\mathbb{R})$, the matrices with nonvanishing determinants.

Let $\| \ \|$ be a norm on V (c.f. [L], [N]). In the usual way there is induced an operator norm, also denoted $\| \ \|$, on $\mathrm{End}(V)$. We recall the definition:

$$\|A\| = \sup\left\{ \frac{\|Av\|}{\|v\|} : v \in V - \{0\} \right\} \quad A \in \mathrm{End}(V). \tag{3.1}$$

The norm on $\mathrm{End}V$ makes $\mathrm{End}(V)$ into a metric space. Since the determinant is a continuous function on $\mathrm{End}(V)$, we know that $\mathrm{GL}(V)$ is an open subset of $\mathrm{End}V$ (see also (3.6) below), so it also is a metric space.

DEFINITION. A *one-parameter group of linear transformations* of V is a continuous homomorphism

$$M : \mathbb{R} \to \mathrm{GL}(V). \tag{3.2}$$

Thus $M(t)$ is a collection of linear maps such that

(i) $M(0) = 1_V$, the identity of V,

(ii) $M(s)M(t) = M(s+t) \qquad s, t \in \mathbb{R}$,

(iii) $M(t)$ depends continuously on t.

REMARKS. (a) The topology on $GL(V)$ is easily verified to be a group topology as defined in 1. Thus, for $A \in \operatorname{End} V$ and $r > 0$, set

$$\mathscr{B}_r(A) = \{A' \in \operatorname{End} V : \|A' - A\| < r\}. \tag{3.3}$$

First, the basic formula [N, p. 76],

$$\|AB\| \leqslant \|A\| \, \|B\| \tag{3.4}$$

implies that left and right multiplication are continuous. Hence the topology is homogeneous. Then the Neumann formula [N, p. 177],

$$(1_V - A)^{-1} = \sum_{n=0}^{\infty} A^n \tag{3.5}$$

valid for A with $\|A\| < 1$ shows that

$$\mathscr{B}_r(1_V))^{-1} \subseteq \mathscr{B}_s(1_V) \tag{3.6}$$

with $s = r/(1 - r)$. Similarly the formula

$$(1_V + A)(1_V + B) = 1_V + A + B + AB \tag{3.7}$$

shows

$$\mathscr{B}_r(1_V)\mathscr{B}_s(1_V) \subseteq \mathscr{B}_{r+s+rs}(1_V).$$

Thus all the conditions of Lemma 3 are checked, and we have a group topology.

(b) Furthermore, it is not difficult to verify that the topology defined by the norm coincides with the compact-open topology defined in §1 on $\mathrm{GL}(V)$ as a subgroup of $\mathrm{Hm}(V)$. This is left as an exercise. Hence this definition of one-parameter group is a special case of the definition of §2.

For $A \in \operatorname{End} V$, define

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}. \tag{3.8}$$

Since $\|A^n\| \leqslant \|A\|^n$, we see, by the standard estimates in the exponential series, that the series defining $\exp A$ converges absolutely for all A and uniformly on any $\mathscr{B}_r(0)$. Hence exp defines a smooth, in fact analytic, map from $\operatorname{End}(V)$ to itself. We will see shortly that in fact $\exp A \in \mathrm{GL}(V)$.

PROPOSITION 8. *If A and B in* End V *commute with each other, then*

$$\exp(A + B) = \exp A \exp B. \tag{3.9}$$

Proof. Computing formally we have

$$\begin{aligned}\exp A \exp B &= \left(\sum_{n=0}^{\infty} \frac{A^n}{n!}\right)\left(\sum_{m=0}^{\infty} \frac{B^m}{m!}\right) = \sum_{n,m=0}^{\infty} \frac{A^n B^m}{n!m!} \\ &= \sum_{l=0}^{\infty} \frac{1}{l!}\left(\sum_{m+n=l} \frac{l!}{m!n!} A^n B^m\right) \\ &= \sum_{l=0}^{\infty} \frac{1}{l!}\left(\sum_{k=0}^{l} \binom{l}{k} A^k B^{l-k}\right).\end{aligned}$$

If A and B commute, the familiar binomial formula applies and says

$$(A + B)^l = \sum_{k=0}^{l} \binom{l}{k} A^k B^{l-k}.$$

Substituting this in our formula for $\exp A \exp B$, and noting that all manipulations are valid because the series converge absolutely, we see the proposition follows. ■

COROLLARY 9. *For any* $A \in \operatorname{End} V$, *the map* $t \to \exp(tA)$ *is a one-parameter group of linear transformations on* V. *In particular* $\exp A \in \mathrm{GL}(V)$ *and* $(\exp(A))^{-1} = \exp(-A)$.

Proof. Since for any real numbers s and t the matrices sA and tA commute with one another, this corollary follows immediately from Proposition 8. ■

The main result of this section is the converse of Corollary 9.

THEOREM 10. *Every one-parameter group* M *of linear transformations of* V *has the form*

$$M(t) = \exp(tA) \tag{3.10}$$

for some $A \in \operatorname{End} V$.

The transformation A is called the *infinitesimal generator* of the group $t \to \exp(tA)$. The flow illustrated in Figure 1 is in fact given by a one-parameter group with infinitesimal generator $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

REMARK. Since for $v \in V$ we have

$$(\exp tA)(v) = v + tA(v) + \sum_{n=2}^{\infty} \frac{t^n A^n(v)}{n!},$$

the infinitesimal generator of the one-parameter group $M(t) = \exp(tA)$ can be computed by the formula

$$A(v) = \lim_{t \to 0} \frac{M(t)(v) - v}{t} = \frac{d}{dt}(M(t)(v))\Big|_{t=0}. \tag{3.11}$$

Thus the one-parameter group $M(t)$ is associated by the discussion at the end of §2 to the system of differential equations

$$\frac{dv}{dt} = A(v). \tag{3.12}$$

These equations are of course basic in the theory of linear systems, which is applied in electrical engineering, economics, etc. If we know that $M(t)(v)$ is differentiable, then the existence and uniqueness theorem for differential equations implies Theorem 10, but we do not know *a priori* that $M(t)$ is differentiable. The burden of the proof of Theorem 3 is to get around this ignorance, thereby establishing that a merely continuous map $t \to M(t)$ satisfying the group law (3.2) (ii) is in fact analytic. This is a recurrent theme in Lie theory, and is also expressed in the main theorem (Theorem 17) of this paper. It found its ultimate expression in Hilbert's 5th Problem: to show that if a topological group is locally (i.e., a neighborhood of every point is) homeomorphic to Euclidean space, then the group is in fact an analytic manifold with analytic group law (a Lie group). This problem was resolved positively in the early 1950's by A. Gleason [**G**]. See also [**Ka**], [**MZ**].

We take up now the proof of Theorem 10. It will require some preliminary results.

Let $\mathscr{B}_r(A)$ be the open ball of radius r around A, as defined in formula (3.3).

PROPOSITION 11. *For sufficiently small* $r > 0$, *the map* exp *takes* $\mathscr{B}_r(0)$ *bijectively onto an open neighborhood of* 1_V *in* $\mathrm{GL}(V)$. *One has* $\exp(\mathscr{B}_r(0)) \subseteq \mathscr{B}_s(1_V)$ *where* $s = e^r - 1$.

Proof. Let $D\exp_A$ be the differential of exp at A. It is a linear map from $\operatorname{End}(V)$ to $\operatorname{End}(V)$ defined by

$$D\exp_A(B) = \lim_{t \to 0} \frac{\exp(A + tB) - \exp A}{t}.$$

From the definition (3.8) of exp, it is easy to compute that

$$D\exp_0(B) = B.$$

That is $D\exp_0$ is the identity map on $\mathrm{End}(V)$. In particular $D\exp_0$ is invertible. Therefore the first statement of the proposition follows from the Inverse Function Theorem [**L**], [**R**]. The inclusion $\exp(\mathscr{B}_r(0)) \subseteq \mathscr{B}_s(1_V)$ follows from the obvious termwise estimation of $\exp(A) - 1_V$. ■

REMARK. If one defines

$$\log(1_V - A) = -\sum_{n=1}^{\infty} \frac{A^n}{n}, \tag{3.13}$$

then just as for real numbers, one sees this series converges absolutely for $\|A\| < 1$. Further, for all $B \in \mathscr{B}_1(1_V)$ one has

$$\exp(\log B) = B. \tag{3.14}$$

Formula (3.14) is known in the scalar case, and this implies that in fact (3.14) is an identity in absolutely convergent power series, whence it follows in the matrix case. The formulas (3.13) and (3.14) allow an alternate proof of Proposition 11 which avoids appeal to the Inverse Function Theorem and gives the explicit estimate that exp is 1-1 on $\mathscr{B}_{\log 2}(0)$. However, this explicit value of r is not needed, and we need in any case to appeal to the Inverse Function Theorem below in Theorem 17, so this more explicit proof of Proposition 11 gives us no particular benefit.

PROPOSITION 12. *Choose an $r < \log 2$, and let T be in $\exp \mathscr{B}_r(0)$, say $T = \exp A$. Then the transformation $S = \exp(A/2)$ is a square root of T; that is, $S^2 = T$. Moreover, S is the unique square root of T contained in $\exp \mathscr{B}_r(0)$.*

Proof. That $S^2 = T$ follows directly from Proposition 8. It is only necessary to prove the uniqueness of S. From Proposition 11, we see that our restriction on r implies $\exp \mathscr{B}_r(0) \subseteq \mathscr{B}_1(1_V)$. Hence it will suffice to show that if A, B are distinct linear maps of norm less than 1, then $(1_V + A)^2 \neq (1_V + B)^2$. Suppose the contrary. Then expanding the squares, cancelling the 1_V's and transposing, we find the equation

$$2(A - B) = B^2 - A^2 = B(B - A) + (B - A)A.$$

Taking norms yields

$$2\|A - B\| \leqslant \|B\|\,\|B - A\| + \|B - A\|\,\|A\| = (\|B\| + \|A\|)\|B - A\|.$$

This implies either $\|A - B\| = 0$, which is false since $A \neq B$, or $\|A\| + \|B\| \geqslant 2$, which is false since both $\|A\|$ and $\|B\|$ are less than 1. This contradiction establishes the uniqueness of S. ■

Proof of Theorem 10. Let $t \to M(t)$ be a continuous one-parameter group in $\mathrm{GL}(V)$. Since $M(0) = 1_V$, if we specify $r > 0$, we may by continuity and Proposition 11 find an $\varepsilon > 0$ such that $M(t) \in \exp(\mathscr{B}_r(0))$ for $|t| \leqslant \varepsilon$. We take $r < \log 2$. Write

$$M(\varepsilon) = \exp A_1$$

for appropriate $A_1 \in \mathscr{B}_r(0)$. If we set

$$A = \left(\frac{1}{\varepsilon}\right) A_1,$$

then $M(\varepsilon) = \exp(\varepsilon A)$. The transformations $M(\varepsilon/2)$ and $\exp((\varepsilon/2)A)$ are then both square roots of $M(t)$ lying in $\exp(\mathscr{B}_r(0))$. By Proposition 12 we conclude

$$M(\varepsilon/2) = \exp((\varepsilon/2)A).$$

An obvious induction using Proposition 12 shows that

$$M(2^{-n}\varepsilon) = \exp(2^{-n}\varepsilon A)$$

for all positive integers n. Taking mth powers, we conclude

$$M(m2^{-n}\varepsilon) = \exp(m2^{-n} \in A)$$

for all integers m and n. Since the numbers $m2^{-n}\varepsilon$ are dense in $\mathbb{R}$, Theorem 10 follows by continuity. ■

4. Properties of the Exponential Map

The map exp is the basic link between the linear structure on End V and the multiplicative structure on $\mathrm{GL}(V)$. We will describe some salient properties of this link.

Choose r with $0 < r \leqslant 1/2$ such that exp is one-to-one on $\mathscr{B}_r(0)$. Choose $r_1 < r$ so that if $A, B \in \mathscr{B}_{r_1}(0)$, then $\exp A \exp B$ is contained in $\exp \mathscr{B}_r(0)$. Then we can write

$$\exp A \exp B = \exp C \tag{4.1}$$

for some $C \in \mathscr{B}_r(0)$. The Inverse Function Theorem guarantees that C is a smooth (in fact analytic) function of A and B. There is a beautiful formula, the Campbell-Hausdorff formula [**J1**], [**Se**], which expresses C as a universal power series in A and B. To develop this completely would take too long. We will just give the first two terms in the expression for C. These suffice for most purposes.

For $A, B \in \mathrm{End}\ V$, write

$$[A, B] = AB - BA. \tag{4.2}$$

The quantity $[A, B]$ is called the *commutator* of A and B, and will be seen later to provide the Lie bracket operation in the Lie algebras we construct.

PROPOSITION 13. *Suppose A, B, C have norm at most $1/2$ and satisfy equation* (4.1). *Then we have*

$$C = A + B + \frac{1}{2}[A, B] + S, \tag{4.3}$$

where the remainder term S satisfies

$$\|S\| \leqslant 65(\|A\| + \|B\|)^3. \tag{4.4}$$

Proof. We have

$$\exp C = 1_V + C + R_1(C), \tag{4.5}$$

where the remainder $R_1(C)$ is

$$R_1(C) = \sum_{n=2}^{\infty} \frac{C^n}{n!}$$

and satisfies the obvious estimate

$$\|R_1(C)\| \leqslant \|C^2\| \left(\sum_{n=2}^{\infty} \frac{\|C\|^{n-2}}{n!} \right) \leqslant \|C\|^2$$

when $\|C\| \leqslant 1$, hence certainly when $\|C\| \leqslant 1/2$.

Similarly we have

$$\exp A \exp B = 1_V + A + B + R_1(A, B), \tag{4.6}$$

where by rearrangement of the double sum

$$R_1(A, B) = \sum_{n=2}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} \right).$$

Hence we have the estimate

$$\|R_1(A, B)\| \leqslant (\|A\| + \|B\|)^2 \left(\sum_{n=2}^{\infty} \frac{(\|A\| + \|B\|)^{n-2}}{n!} \right) \leqslant (\|A\| + \|B\|)^2$$

when $\|A\| + \|B\| \leqslant 1$.

Comparing equations (4.5) and (4.6), we see that equation (4.1) implies

$$C = A + B + R_1(A, B) - R_1(C). \tag{4.7}$$

Hence

$$\|C\| \leqslant \|A\| + \|B\| + (\|A\| + \|B\|)^2 + \|C\|^2 \leqslant 2(\|A\| + \|B\|) + \frac{1}{2}\|C\|$$

when A, B and C all have norm at most $\frac{1}{2}$.

Thus

$$\|C\| \leqslant 4(\|A\| + \|B\|). \tag{4.8}$$

Returning to equation (4.7), we further find

$$\begin{aligned} \|C - (A + B)\| &\leqslant \|R_1(A, B)\| + \|R_1(C)\| \leqslant (\|A\| + \|B\|)^2 + (4(\|A\| + \|B\|))^2 \\ &= 17(\|A\| + \|B\|)^2. \end{aligned} \tag{4.9}$$

We now refine these estimates to second order. In analogy with (4.5) we have

$$\exp C = 1_V + C + \frac{C^2}{2} + R_2(C), \tag{4.10}$$

where

$$R_2(C) = \sum_{n=3}^{\infty} \frac{C^n}{n!}$$

is easily estimated by

$$\|R_2(C)\| \leqslant \left(\frac{1}{3}\right)\|C\|^3 \tag{4.11}$$

when $\|C\| \leqslant 1$.

If we substitute expression (4.3) for C in equation (4.10), we obtain

$$\begin{aligned} \exp C &= 1_V + A + B + \frac{1}{2}[A, B] + S + \frac{1}{2}C^2 + R_2(C) \\ &= 1_V + A + B + \frac{1}{2}[A, B] + \frac{1}{2}(A + B)^2 + T \\ &= 1_V + A + B + \frac{1}{2}(A^2 + 2AB + B^2) + T, \end{aligned} \tag{4.12}$$

where

$$T = S + \frac{1}{2}\left(C^2 - (A + B)^2\right) + R_2(C).$$

On the other hand, we have

$$\exp A \exp B = 1_V + A + B + \frac{1}{2}(A^2 + 2AB + B^2) + R_2(A, B), \tag{4.13}$$

where

$$R_2(A, B) = \sum_{n=3}^{\infty} \frac{1}{n!}\left(\sum_{k=0}^{\infty} \binom{n}{k} A^k B^{n-k}\right)$$

satisfies $\|R_2(A, B)\| \leqslant \frac{1}{3}(\|A\| + \|B\|)^3$ when $\|A\| + \|B\| \leqslant 1$.

Comparison of (4.12) and (4.13) in the light of (4.1) yields

$$S = R_2(A, B) + \frac{1}{2}\left((A + B)^2 - C^2\right) - R_2(C).$$

Taking norms, we find

$$\begin{aligned}\|S\| &\leqslant \|R_2(A,B)\| + \frac{1}{2}(\|(A+B)(A+B-C)+(A+B-C)C\|) + \|R_2(C)\| \\ &\leqslant \frac{1}{3}(\|A\|+\|B\|)^3 + \frac{1}{2}(\|A\|+\|B\|+\|C\|)\|A+B-C\| + \frac{1}{3}\|C\|^3 \\ &\leqslant \frac{1}{3}(\|A\|+\|B\|)^3 + \frac{5}{2}(\|A\|+\|B\|)\cdot 17(\|A\|+\|B\|)^2 + \frac{1}{3}(4(\|A\|+\|B\|))^3 \\ &\leqslant 65(\|A\|+\|B\|)^3,\end{aligned}$$

as was to be shown.

We will derive two main consequences of Proposition 13. These relate group operations in $\mathrm{GL}(V)$ to the linear operations in $\mathrm{End}(V)$, and are crucial ingredients in the proof of the main theorem (Theorem 17 in §5) that relates Lie algebras to Lie groups. Proposition 14 relates group multiplication in $\mathrm{GL}(V)$ to addition in $\mathrm{End}(V)$, and Proposition 15 relates the group commutator operation to the bilinear commutator bracket defined in equation (4.2).

PROPOSITION 14 (Trotter Product Formula). *For $A, B \in \mathrm{End}\, V$, one has*

$$\exp(A+B) = \lim_{n\to\infty} (\exp(A/n)\exp(B/n))^n. \tag{4.14}$$

Proof. For n large enough, A/n and B/n will be close enough to the origin that formula (4.3) applies. We then have

$$\exp(A/n)\exp(B/n) = \exp C_n,$$

where by estimate (4.9)

$$\|C_n - (A+B)/n\| \leqslant 17((\|A\|+\|B\|)/n)^2.$$

Hence as $n \to \infty$, we see that $nC_n \to A+B$. Since $\exp nC_n = (\exp C_n)^n$, equation (4.14) follows. ■

Recall that the (linear) commutator $[A, B]$ is defined in equation (4.2). Recall also that if g, h are elements of a group, then the group commutator of g and h, written $(g:h)$, is the expression

$$(g:h) = ghg^{-1}h^{-1}.$$

PROPOSITION 15 (Commutator formula). *For $A, B \in \mathrm{End}\, V$, one has*

$$\begin{aligned}\exp[A,B] &= \lim_{n\to\infty} (\exp(A/n)\exp(B/n)\exp(-A/n)\exp(-B/n))^{n^2} \\ &= \lim_{n\to\infty} ((\exp(A/n):\exp(B/n))^{n^2}.\end{aligned} \tag{4.15}$$

Proof. As in Proposition 14, for large n we have

$$\exp(A/n)\exp(B/n) = \exp C_n = \exp\left((A+B)/n + \frac{1}{2}\frac{[A,B]}{n^2} + S_n\right),$$

where

$$\|S_n\| \leqslant 65\frac{(\|A\|+\|B\|)^3}{n^3}.$$

Similarly

$$\exp(-A/n)\exp(-B/n) = \exp\left(-(A+B)/n + \left(\frac{1}{2}\right)\frac{[A,B]}{n^2} + S_n'\right) = \exp C_n'$$

with also

$$\|S_n'\| \leqslant 65\frac{(\|A\| + \|B\|)^3}{n^3}.$$

Hence

$$(\exp(A/n) : \exp(B/n)) = \operatorname{Exp} C_n \exp C_n' = \operatorname{Exp} E_n,$$

where

$$\begin{aligned} E_n &= C_n + C_n' + \frac{1}{2}[C_n, C_n'] + T_n \\ &= \frac{[A, B]}{n^2} + \frac{1}{2}[C_n, C_n'] + S_n + S_n' + T_n, \end{aligned}$$

where T_n is the term S in equation (4.3) if $A = C_n$ and $B = C_n'$.

It will suffice to show that there is a number γ, depending on A and B, such that

$$\left\| E_n - \frac{[A, B]}{n^2} \right\| \leqslant \frac{\gamma}{n^3}.$$

For then

$$(\exp E_n)^{n^2} = \exp([A, B] + U_n)$$

with $\|U_n\| \leqslant \gamma/n$, and equation (4.15) follows. In turn, it will suffice to show that the 2nd, 3rd, 4th and 5th terms in the expression for E_n are each less than a constant times n^{-3}. For S_n, S_n' and T_n, this follows from Proposition 13. Thus we need only worry about $[C_n, C_n']$. We compute

$$\begin{aligned} [C_n, C_n'] &= \left[\frac{1}{n}(A + B) + \frac{1}{2n^2}[A, B] + S_n, \frac{-1}{n}(A + B) + \frac{1}{2n^2}[A, B] + S_n'\right] \\ &= \frac{1}{n^3}[A + B, [A, B]] + \frac{1}{n}[A + B, S_n + S_n'] + \frac{1}{2n^2}[[A, B], S_n' - S_n] \\ &\quad + [S_n, S_n']. \end{aligned}$$

Using Proposition 13, we see that each of the four terms in this last sum is bounded by a constant times n^{-3}. (In fact, all terms except the first are bounded by a constant times n^{-4}).

There is one further concept involving the exponential map that is basic to Lie theory. It involves conjugation, which is generally referred to as the "adjoint action." For $g \in \mathrm{GL}(V)$ and $T \in \operatorname{End} V$, we can form the conjugate

$$\operatorname{Ad} g(A) = gAg^{-1}. \tag{4.16}$$

The following proposition is easily verified and left as an exercise.

PROPOSITION 16. (i) $\operatorname{Ad} g(aA + bB) = a \operatorname{Ad} g(A) + b \operatorname{Ad} g(B)$ *for* $A, B \in \operatorname{End} V$; $a, b \in \mathbb{R}$; *and* $g \in \mathrm{GL}(V)$.

(ii) $\operatorname{Ad} g(AB) = \operatorname{Ad} g(A) \operatorname{Ad} g(B)$.

(iii) $\operatorname{Ad} g_1 g_2(A) = \operatorname{Ad} g_1(\operatorname{Ad} g_2(A))$.

Formulas (i) and (ii) say $\operatorname{Ad} g$ is an algebra automorphism of $\operatorname{End} V$, and Formula (iii) says the map $\operatorname{Ad}: g \to \operatorname{Ad} g$ is a group homomorphism from $\mathrm{GL}(V)$ to the automorphism group of $\operatorname{End}(V)$. The map Ad is called the *adjoint action* of $\mathrm{GL}(V)$ on $\operatorname{End}(V)$.

Formula (iii) implies in particular that if $\exp tA$ is a one-parameter subgroup of $\mathrm{GL}(V)$, then $\operatorname{Ad} \exp tA$ is a one-parameter group of linear transformations on $\operatorname{End} V$. Hence $\operatorname{Ad} \exp tA$ has infinitesimal generator $\mathscr{A} \in \operatorname{End}(\operatorname{End} V)$. We can compute $\mathscr{A}$ by the formula

$$\begin{aligned}\mathscr{A}(B) &= \lim_{t\to 0}\frac{(\exp tA)B(\exp(-tA)) - B}{t}\\ &= \frac{d}{dt}(\exp tA)B(\exp(-tA))|_{t=0}\\ &= (A(\exp tA)B(\exp - tA) + (\exp tA)B(-A)(\exp - tA))|_{t=0}\\ &= AB - BA = [A, B].\end{aligned}$$

Here we have used the fact that

$$\frac{d}{dt}(\exp(tA)) = A\exp(tA).$$

This formula may be verified by direct calculation from the definition of $\exp(tA)$. Hence if we define

$$\operatorname{ad}A : \operatorname{End} V \to \operatorname{End} V$$

by

$$\operatorname{ad}A(B) = [A, B],$$

we have the following formula.

PROPOSITION 17. *For* $A \in \operatorname{End} V$

$$\operatorname{Ad}(\exp A) = \exp(\operatorname{ad}A). \tag{4.17}$$

5. The Lie Algebra of a Matrix Group

By a *matrix group* we mean a closed subgroup of $\mathrm{GL}(V)$ for some vector space V. This section shows a matrix group is a Lie group. What that means is expressed in Theorem 17. Most, though not all, Lie groups can be realized as matrix groups. This article discusses only matrix groups.

EXAMPLES. (i) $\mathrm{GL}(V)$ itself.

(ii) $\mathrm{SL}_n(\mathbb{R})$, the special linear group, of $n \times n$ matrices of determinant 1.

(iii) $O_{p,q}$, the "pseudo-orthogonal groups," consisting of all matrices in $\mathrm{GL}_{p+q}(\mathbb{R})$ that preserve the indefinite inner product

$$(x, x')_{p,q} = \sum_{i=1}^{P} x_i x_i' - \sum_{i=p+1}^{p+q} x_i x_i', \; x, x' \in \mathbb{R}^{p+q}.$$

(iv) $\mathrm{SP}_{2n}(\mathbb{R})$, the real symplectic group, consisting of all matrices in $\mathrm{SL}_{2n}(\mathbb{R})$ that preserve the skew-symmetric bilinear form

$$\langle x, x'\rangle = \sum_{i=1}^{n} x_i x_{i+n}' - x_i' x_{i+n} x, \; x' \in \mathbb{R}^{2n}.$$

(v) The group $P(U)$ of transformations that preserve a subspace U of V. For instance, if $V = \mathbb{R}^n$, and $U_m = \mathbb{R}^m = \{(x_1, x_2, \ldots, x_m, 0, 0, \ldots, 0)\}$, where $m \leqslant n$, then

$$P(U_m) = \left\{\begin{bmatrix} A & X \\ 0 & B\end{bmatrix} : A \in \mathrm{GL}_m(\mathbb{R}), B \in \mathrm{GL}_{n-m}(\mathbb{R}), X \in M_{m,n-m}(\mathbb{R})\right\}.$$

Here $M_{m,n-m}(\mathbb{R})$ is the space of $m \times (n-m)$ real matrices.

(vi) Any intersection of matrix groups is a matrix group. For instance, the intersection $\cap_{m=1}^{n} P(U_m)$ of the groups $P(U_m)$ of example (v) is the group of invertible upper triangular matrices.

(vii) The group preserving some closed subgroup, not necessarily a subspace, of V. For example, let $\mathbb{Z}^n \subseteq \mathbb{R}^n$ be the discrete subgroup of vectors with integral entries. Set

$$\mathrm{GL}_n(\mathbb{Z}) = \{A \in \mathrm{GL}_n(\mathbb{R}) : A(\mathbb{Z}^n) = \mathbb{Z}^n\}.$$

Then $\mathrm{GL}_n(\mathbb{Z})$ can also be shown to consist of matrices with integer entries and determinant ± 1.

(viii) The group commuting with some family $\{T_i\}$ of operators on V is a matrix group. For example, we can identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ by letting x_{2j-1} and x_{2j} be the real and imaginary parts of the coordinate z_j of $z = (z_1, z_2, \ldots, z_n) \in \mathbb{C}^n$. If we do so, the operation of multiplication by a complex scalar becomes some (real) linear operator on $\mathbb{R}^{2n}$. Further, the group $\mathrm{GL}_n(\mathbb{C})$ becomes identified with the subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$ formed by elements which commute with the multiplications by complex scalars.

(ix) If G_i is a matrix group in $\mathrm{GL}(V_i)$, $i = 1, 2$, then $G_1 \times G_2$ is a matrix group in $\mathrm{GL}(V_1 \oplus V_2)$ in the obvious way.

(x) If G is a matrix group, then G^0, the connected component of the identity in G, is a matrix group.

(xi) The normalizer in $\mathrm{GL}(V)$ of a matrix group is a matrix group.

The main result of this section is the essential phenomenon behind Lie theory: a matrix group has naturally attached to it a Lie algebra. Before showing this we recall what a Lie algebra is.

DEFINITION. A real *Lie algebra* $\mathfrak{g}$ is a real vector space equipped with a product

$$[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \tag{5.1}$$

satisfying the identities

(5.2)

(i) (Bilinearity). For $a, b \in \mathbb{R}$ and $x, y, z \in \mathfrak{g}$,

$$[ax + by, z] = a[x, z] + b[y, z]$$
$$[z, ax + by] = a[z, x] + b[z, y].$$

(ii) (Skew symmetry). For $x, y \in \mathfrak{g}$,

$$[x, y] = -[y, x].$$

(iii) (Jacobi Identity). For $x, y, z \in \mathfrak{g}$,

$$[x,[y, z]] + [z,[x, y]] + [y,[z, x]] = 0.$$

The first main example of a Lie algebra is End V equipped with the bracket operation $[\,,]$ of commutator, as given in equation (4.2). It is left as an exercise to verify that this satisfies the correct identities. Any subspace of End V which is closed under $[\,,]$ will become a Lie algebra in its own right. Since our main theorem will provide us with such a subspace for each matrix group, we will postpone a more explicit discussion of examples.

Consider a matrix group $G \subseteq \mathrm{GL}(V)$. Let $\exp^{-1}(G) \subseteq \mathrm{End}\, V$ be the inverse image of G under exp. Since $\exp(nA) = (\exp A)^n$, it is clear that $\exp^{-1}(G)$ is closed under scalar multiplication by integers. Set

$$\mathfrak{g} = \{A \in \mathrm{End}\, V : \exp tA \in G \text{ for all } t \in \mathbb{R}\} = \bigcap_{t \in \mathbb{R}^x} t \exp^{-1}(G).$$

Observe that $\mathfrak{g}$ is the collection of infinitesimal generators of one-parameter subgroups of G. We call $\mathfrak{g}$ the *Lie algebra* of G.

THEOREM 17. (a) *The Lie algebra* $\mathfrak{g}$ *of a matrix group* G *is a Lie algebra.*

(b) *The map* $\exp : \mathfrak{g} \to G$ *maps a neighborhood of* 0 *in* $\mathfrak{g}$ *bijectively onto a neighborhood of* 1_V *in* G.

REMARKS. (i) Part (b) of Theorem 17 implies G is locally homeomorphic to Euclidean space. In

fact it is not hard to refine part (b) and show that G has the structure of a smooth manifold, such that the group multiplication is smooth, but we will not do that here.

(ii) Theorem 17 provides a geometric picture of the relation between $\mathfrak{g}$ and G. If a one-parameter group $\exp(tA)$ is regarded as a curve inside the vector space $\operatorname{End} V$, then this curve passes through the identity 1_V at time $t = 0$. By differentiating the formula for $\exp tA$, we see the tangent vector at the point 1_V to this curve is just A. Thus, as we have defined it, $\mathfrak{g}$ consists simply of all tangent vectors to the curves defined by one-parameter groups in G. But Theorem 17 asserts that these tangent vectors actually fill out some linear subspace (namely $\mathfrak{g}$) of $\operatorname{End} V$, and further, if we make the smooth change of coordinates $A \to \exp A$, then this linear subspace $\mathfrak{g}$ is bent in such a way that it lies entirely in G, and fills up G around 1_V. In other words, G is shown to be a smooth multidimensional surface inside $\operatorname{End} V$, and $\mathfrak{g}$ is simply its tangent space at the point 1_V.

The main burden of the proof of Theorem 17 is carried by the following technical result.

LEMMA 18. *Suppose $\{A_n\}$ is a sequence in $\exp^{-1}(G)$, and $\|A_n\| \to 0$. Let s_n be a sequence of real numbers. Then any cluster point of s_nA_n is in $\mathfrak{g}$.*

Proof. Let B be the cluster point. By passing to a subsequence if necessary we may assume that s_nA_n converges to B. Fix a number $t \in \mathbb{R}$. Let m_n be an integer such that $|m_n - ts_n| \leqslant 1$. Then m_nA_n converges to tB; for we have

$$\begin{aligned}\|m_nA_n - tB\| &= \|(m_n - ts_n)A_n + t(s_nA_n - B)\| \\ &\leqslant |m_n - ts_n|\,\|A_n\| + |t|\,\|s_nA_n - B\| \\ &\leqslant \|A_n\| + |t|\,\|s_nA_n - B\|\end{aligned}$$

which converges to zero as $n \to \infty$, by our assumptions on A_n and B. Since $m_nA_n \in \exp^{-1}(G)$, and $\exp^{-1}(G)$ is closed, we see that $tB \in \exp^{-1}(G)$. Since t was arbitrary in $\mathbb{R}$, we see that $B \in \mathfrak{g}$. ■

Proof of Theorem 17. We first show $\mathfrak{g}$ is a subspace of $\operatorname{End} V$. Since $\mathfrak{g}$ is by definition closed under scalar multiplication, we need only show it is closed under addition. Take $A, B \in \mathfrak{g}$. Then as in Proposition 14 we know that for large enough n

$$\exp(A/n)\exp(B/n) = \exp C_n,$$

where $\|C_n\| \to 0$, and $nC_n \to A + B$. Hence Lemma 18 implies $A + B \in \mathfrak{g}$.

Next we show that if $A, B \in \mathfrak{g}$, then also $[A, B] \in \mathfrak{g}$. As in Proposition 15 we know that for large n we have

$$(\exp(A/n) : \exp(B/n)) = \exp E_n$$

with $E_n \to 0$ and $n^2E_n \to [A, B]$. Another application of Lemma 18 says $[A, B] \in \mathfrak{g}$. This concludes part (a) of Theorem 17.

We know $\mathfrak{g}$ is a linear subspace of $\operatorname{End}(V)$. Let $Y \subseteq \operatorname{End}(V)$ be a complementary subspace of $\mathfrak{g}$, so that $\operatorname{End} V = \mathfrak{g} \oplus Y$. Let p_1 and p_2 be the projections of $\operatorname{End} V$ on $\mathfrak{g}$ and Y, respectively, with respective kernels Y and $\mathfrak{g}$. Define a map $E : \operatorname{End} V \to \operatorname{GL}(V)$ by

$$E(A) = \exp(p_1(A))\exp(p_2(A)).$$

By use of Proposition 13, we can compute that

$$\frac{d}{dt}\left(\exp(p_1(tA))\exp(p_2(tA))\right)\Big|_{t=0} = p_1(A) + p_2(A) = A.$$

This says that the differential of E at 0 is the identity map on $\operatorname{End} V$, so that E takes small neighborhoods of 0 to neighborhoods of 1_V bijectively, by the Inverse Function Theorem. Choose a small ball $\mathscr{B}_r(0) \subseteq \operatorname{End} V$, and suppose $\exp(\mathscr{B}_r(0) \cap \mathfrak{g})$ does not cover a neighborhood of 1_V in G. Then we can find a sequence $B_n \in \exp^{-1}(G)$ such that $B_n \to 0$, but $B_n \notin \mathfrak{g}$. When B_n is close enough to 0, we may write

$$\exp B_n = E(A_n)$$

for some A_n. We will have $A_n \to 0$ as $B_n \to 0$. Then

$$\exp(p_2(A_n)) = \exp(p_1(A_n))^{-1} \exp B_n$$

is also in G, and is nonzero by our assumption on B_n. Since $A_n \to 0$, $p_2(A_n) \to 0$ also. The sequence $\|p_2(A_n)\|^{-1} p_2(A_n)$ will have cluster points, and these must be in $\mathfrak{g}$ by Lemma 18. On the other hand, $p_2(A_n) \in Y$, so all cluster points must be in Y. This contradicts the fact that Y was chosen complementary to $\mathfrak{g}$, so statement (b) of Theorem 17 follows. ∎

EXAMPLES. We will describe below the Lie algebras of some of the groups listed at the beginning of this section. The verification that the indicated Lie algebras are indeed the Lie algebras of the stated groups is left as an exercise.

(i) The Lie algebra of $\mathrm{GL}(V)$ is of course $\mathrm{End}(V)$.
(ii) The Lie algebra of $\mathrm{SL}_n(\mathbb{R})$ is the space of $\mathrm{sl}_n(\mathbb{R})$ of $n \times n$ matrices of trace zero.
(iii) Let β be a bilinear form on V. The *isometry group of* β is the group of invertible operators A such that

$$\beta(Au, Av) = \beta(u, v) \qquad \text{for all} \qquad u, v \in V.$$

The Lie algebra of this group is the space of operators B such that

$$\beta(Bu, v) + \beta(u, Bv) = 0.$$

In particular the Lie algebra $\mathfrak{o}_n(V)$ of the orthogonal group $O_n(\mathbb{R})$ of isometries of the standard inner product on $\mathbb{R}^n$ is the space of skew-symmetric matrices.
(iv) The Lie algebra of the subgroup of $\mathrm{GL}(V)$ of maps commuting with given operators $\{T_i\}$ is the subalgebra of $\mathrm{End}\, V$ commuting with the T_i.
(v) The Lie algebra of the group $P(\{V_i\})$ of invertible transformations which preserve each of the subspaces V_i of V is the subalgebra of all transformations which preserve the V_i. In particular, the Lie algebra of the group of invertible upper triangular matrices is the vector space of all upper triangular matrices.
(vi) The Lie algebra of $G_1 \cap G_2$, for matrix groups G_i, is $\mathfrak{g}_1 \cap \mathfrak{g}_2$.
(vii) A matrix group G and its identity component G^0 have the same Lie algebra.

After its existence, the second most important feature of $\mathfrak{g}$ is that it is natural (in the sense of category theory). This is the content of our next theorem.

Let $\mathfrak{g}$, $\mathfrak{h}$, be real Lie algebras. A *homomorphism* from $\mathfrak{g}$ to $\mathfrak{h}$ is a linear map

$$L : \mathfrak{g} \to \mathfrak{h}$$

satisfying

$$L([x, y]) = [Lx, Ly] \qquad x, y \in \mathfrak{g}. \tag{5.3}$$

Let V, U be real vector spaces.

THEOREM 19. *Let $G \subseteq \mathrm{GL}(V)$ be a matrix group with Lie algebra $\mathfrak{g}$. Let $\phi : G \to \mathrm{GL}(U)$ be a continuous homomorphism. Then there is a homomorphism of Lie algebras*

$$d\phi : \mathfrak{g} \to \mathrm{End}\, U \tag{5.4}$$

such that

$$\exp(d\phi(A)) = \phi(\exp A). \tag{5.5}$$

Proof. If $A \in \mathfrak{g}$, then $\exp tA$ is a one-parameter subgroup of G, so $\phi(\exp(tA))$ is a one-parameter subgroup of $\phi(G) \subseteq \mathrm{GL}(U)$. Hence by Theorem 10 we may write $\phi(\exp(tA)) = \exp(tB)$ for some $B \in \mathrm{End}\, U$. If we define

$$d\phi(A) = B,$$

then equation (5.5) will obviously be satisfied. To prove this theorem, it suffices to show that $d\phi$ is a homomorphism of Lie algebras. But this follows directly from Propositions 14 and 15 which show that the Lie algebra operations in $\mathfrak{g}$ are determined by operations in G.

EXAMPLE. The formula (4.17) shows that $\mathfrak{g}$ is an invariant subspace of $\operatorname{End} V$ under the operators $\operatorname{Ad} g$, $g \in G$. The restriction of $\operatorname{Ad} g$ to $\mathfrak{g}$ is again denoted by $\operatorname{Ad} g$, and the resulting action of G on $\mathfrak{g}$ is still called the adjoint action. In terms of Theorem 19, the formula (4.17) has the interpretation

$$d(\operatorname{Ad}) = \operatorname{ad}. \tag{5.6}$$

An immediate consequence of Theorem 19 is:

COROLLARY 20. *If $G_1 \subseteq \mathrm{GL}(V)$ and $G_2 \subseteq \mathrm{GL}(U)$ are isomorphic matrix groups, then their Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are isomorphic as Lie algebras.*

Proof. Let $\phi : G_1 \to G_2$ be a continuous isomorphism with continuous inverse ϕ^{-1}. Then in particular ϕ is a continuous homomorphism from G_1 to $\mathrm{GL}(U)$, and ϕ^{-1} is a continuous homomorphism from G_2 to $\mathrm{GL}(V)$. Theorem 19 therefore provides us with associated Lie algebra homomorphisms $d\phi$ and $d(\phi^{-1})$. It follows from the definition of the Lie algebra of a matrix group and formula (5.5) that in fact $d\phi(\mathfrak{g}_1) \subseteq \mathfrak{g}_2$, and similarly $d(\phi^{-1})(\mathfrak{g}_2) \subseteq \mathfrak{g}_1$. It further follows from formula (5.5) that since $\phi^{-1} \circ \phi$ is the identity on G_1, then also $d(\phi^{-1}) \circ d\phi$ is the identity on $\mathfrak{g}_1$. In other words $d(\phi^{-1}) = (d\phi)^{-1}$, so $d\phi$ is in fact a Lie algebra isomorphism from $\mathfrak{g}_1$ to $\mathfrak{g}_2$. ■

The converse of Corollary 20, that groups with isomorphic Lie algebras are isomorphic, is false. For example the rotation group

$$\mathrm{SO}_2 = \left\{ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

and the diagonal group

$$D_1 = \left\{ \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} : a \in \mathbb{R}, \quad a > 0 \right\}$$

both have Lie algebra isomorphic to $\mathbb{R}$, but SO_2 is homeomorphic to a circle, while D_1 is homeomorphic to $\mathbb{R}$, so they are certainly not isomorphic.

However, the converse of Corollary 20 is in a sense almost true, so that the bracket operation on $\mathfrak{g}$ almost determines G as a group. After the existence of the Lie algebra, this fact is the most remarkable in Lie theory. Its precise formulation is known as Lie's Third Theorem. It is in proving a suitable version of Lie's Third Theorem that Lie theory begins to get involved, so we will leave the story here. Precise treatments of these issues can be found in [**A**], [**Ch**], [**He**], [**Se**].

6. Loose Ends and Further Developments

In §§3, 4, and 5 we have shown that to each matrix group G, there is associated in a close and natural way, a Lie algebra $\mathfrak{g}$, the two being connected via one parameter groups and the exponential map. These facts constitute an important part of the foundations of Lie theory. We will describe briefly what we have omitted from the standard account.

First, we have not treated Lie groups as abstract things-in-themselves, but have only dealt with them as subgroups of a standard group, $\mathrm{GL}(V)$. We could not have discussed abstract Lie groups without assuming the standard language of differentiable manifolds. Our approach allowed us to bring to the fore the remarkable Theorem 17, which asserts that merely the requirements of being closed and being a group inside $\mathrm{GL}(V)$ (or any Lie group) suffices to make the group a smooth manifold. This indicates what a strong regularity condition the group property is. Research over the past decades have continued to underscore this theme [**BT**], [**Ma**], [**Mo**].

Second, we have not demonstrated how complete and mutual is the relationship between Lie groups G and their Lie algebras $\mathfrak{g}$. It is in this direction that the principal technical complications of the theory lie. For example, although we have shown how to attach a Lie algebra to every matrix group, we have not tried to attach a group to every Lie subalgebra of End V. Indeed, this is not possible if one sticks to matrix groups; the one parameter groups obtained by exponentiating elements in a given Lie algebra $\mathfrak{g}$ will generate a group which in a suitable sense has $\mathfrak{g}$ as its Lie algebra but this group will not always be closed in $\mathrm{GL}(V)$. The simplest example is probably the one-parameter group $\exp tA_x$ in $\mathrm{GL}_4(\mathbb{R})$, where

$$A_x = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x \\ 0 & 0 & -x & 0 \end{bmatrix}$$

and x is any irrational number. Also the question of the relation of two matrix groups which have isomorphic Lie algebras, essentially the question of the converse of Corollary 20, involves the notion of covering space and fundamental group [**Ms**] and is beyond the scope of this discussion. Interestingly enough, both these questions are most vexed for the most simple-minded case: abelian Lie groups and their Lie algebras. We close these brief remarks by pointing out that, when G is fairly nonabelian, especially if the center of G is discrete, the existence of the adjoint action and formula (5.6) in particular go a long way toward showing that G is nearly determined by $\mathfrak{g}$. After the foundations comes the rather extensive development of the structure theory of Lie algebras, with direct consequences for the groups. Several fine accounts of the theory of Lie algebras are available, for example [**J**], [**Hu**]. Beyond the theory of Lie groups and algebras in themselves lies the vast domain of their applications. We have mentioned a few of these in the introduction and in §7. Some representative references for applications are [**BC**], [**HP**], [**Hr**], [**Ko**], [**Lo**].

Our treatment in §§3, 4, 5 has been concrete in that we worked only inside End V, but it was also abstract in that it was coordinate free. We record here some common terminology used when bases are introduced. Let $\mathfrak{g} \subseteq \mathrm{End}\, V$ be a Lie subalgebra. Let $\{y_i\}, 1 \leqslant i \leqslant \dim \mathfrak{g}$ be a basis for $\mathfrak{g}$. Then the fact that $\mathfrak{g}$ is a Lie algebra amounts to the statement that the commutators $[y_i, y_j]$ are again linear combinations of the y_k's. Thus we have equations

$$[y_i, y_j] = \sum c_{ij}^k y_k, \tag{6.1}$$

where the c_{ij}^k are real numbers. The equations (6.1) are called the *commutation relations* of the y_i's and the c_{ij}^k are called the *structure constants* of $\mathfrak{g}$ with respect to the y_i.

For example, set

$$e^+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad e^- = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The matrices e^+, e^-, and h form a basis for sl_2, the 2×2 traceless matrices, one of the most fundamental Lie algebras. It is easy to compute that their commutation relations are

$$[h, e^+] = 2e^+ [h, e^-] = -2e^- [e^+, e^-] = h.$$

7. Relations with the Standard Curriculum

In this section we give some examples of how Lie theory makes contact with current staples of undergraduate mathematics. We must of course be very restrictive and brief.

1. Many of the standard theorems of linear algebra are of course also part of the fabric of Lie theory, and gain coherence when considered in that light. For example, several of the standard canonical forms, e.g., Jordan form, the diagonalization of the (skew) Hermitian matrices, amount to classification of the conjugacy classes (orbits under the adjoint action) in a Lie algebra. Jordan form describes conjugacy classes in $\mathrm{End}(\mathbb{C}^n) \sim \mathfrak{gl}_n(\mathbb{C})$, and diagonalization of Hermitian matrices

describes conjugacy classes in U_n, the $n \times n$ unitary group. We cannot explain this interpretation of these results in detail, but encourage the reader to explore it by further reading.

Also, $\mathrm{SL}_n(\mathbb{R})$ or $\mathrm{SL}_n(\mathbb{C})$ are examples of an extremely important class of Lie groups called semisimple groups, and several well-known results in linear algebra are special cases for $\mathrm{SL}_n(\mathbb{R})$ or $\mathrm{SL}_n(\mathbb{C})$ of structure theorems for semisimple groups. (Since GL_n and SL_n are so similar, we state the results for GL_n.) The polar decomposition or singular value decomposition [**St**] says that any $A \in \mathrm{GL}_n(\mathbb{R})$ may be written in the form

$$A = OS = O_1 D O_2,$$

where O, O_1, and O_2 are orthogonal matrices, S is symmetric, and D is diagonal with positive entries. This is the specialization to $\mathrm{GL}_n(\mathbb{R})$ of what is known as the *Cartan decomposition* [**He**] in the context of semisimple Lie groups. Also, the Gram-Schmidt orthonormalization procedure [**St**] says, in group-theoretical terms, that any $A \in \mathrm{GL}_n(\mathbb{R})$ may be written in the form

$$A = OB = ODU,$$

where O is orthogonal, B is upper triangular, D is diagonal with positive entries, and U is upper triangular with diagonal entries all equal to 1. For general semisimple groups, this is known as the *Iwasawa decomposition* [**He**].

Various basic features in the elimination theory, including the "LU factorization" [**St**] of a generic matrix into the product of an upper triangular and a lower triangular matrix, and the "reduced row-echelon form" [**DN**] are aspects of a different kind of decomposition of semisimple groups, known as the *Bruhat decomposition* [**Bo**].

2. The cross product on $\mathbb{R}^3$ defines a Lie algebra structure on $\mathbb{R}^3$. This is in fact isomorphic to $\mathfrak{o}_3$, the Lie algebra of O_3, the 3×3 skew symmetric matrices. The isomorphism is accomplished by

$$(x, y, z) \to \begin{bmatrix} 0 & -x & -y \\ x & 0 & -z \\ y & z & 0 \end{bmatrix}.$$

The generalization of this correspondence to higher dimensions leads to the theory of spinors and Clifford algebras [**J2**].

3. The fact that second mixed partial derivatives are equal is a reflection of the fact that $\mathbb{R}^n$ is an abelian Lie group.

4. The theory of Fourier series and Fourier transform is best understood group-theoretically. See [**Gr**] for a discussion.

5. It is fairly routine in quantum mechanics courses, in conjunction with the Schrödinger equation for the hydrogen atom and angular momentum, to introduce, "raising and lowering operators" [**Me**]. The operators belong to the complexification of $\mathfrak{o}_3$, which is isomorphic to $\mathrm{sl}_2(\mathbb{C})$. The commutation relations of the Lie algebra figure importantly in the computations. The harmonic oscillator is also susceptible to a Lie-theoretic treatment. The Canonical Commutation Relations themselves are the laws for a bracket relation on a Lie algebra, known as the Heisenberg Lie algebra [**Ca**], [**Ho**]. The relations of this algebra with quantum mechanics, and physics generally, is deep and extensive.

6. Perhaps the part of standard undergraduate mathematics that is pedagogically most compatible with Lie theory is differential equations. We have already discussed in §2 how the notion of one-parameter group is a geometrization of the solution of a system of differential equations. And in §3 we noted that one-parameter groups of linear transformations were associated with the very important class of linear, constant coefficient systems. Indeed, the exponential map and linear algebra techniques are often explicitly used in treating these systems [**Br**].

Many of the important classical differential equations are related with Lie theory. Indeed much of the theory of special functions may be considered a branch of Lie theory [Mi], [V]. Below I state, always by way of example, some exercises which I have given to students in differential equations courses and which were favorably received.

A.(i) Let P, Q and the identity operator I span a Lie algebra, with commutation relations $[P, Q] = I$, and of course $[P, I] = [Q, I] = 0$. (These are the Canonical Commutation Relations.) Define $L = (P - I)QP$, and $A_n = (P - I)^n Q^n$ (so $L = A_1 P$). Show that

$$\begin{aligned}
&(a) \quad [Q, (P - I)^n] = -n(P - I)^{n-1}, \\
&(b) \quad A_{n+1} = (A_1 + n)A_n, \\
&(c) \quad [L, A_1] = L - A_1, \text{ and} \\
&(d) \quad L(A_1 + n) = (A_1 + n)L + (L + n) - (A_1 + n).
\end{aligned}$$

(ii) Suppose v_n is an eigenvector for L, with eigenvalue $-n$, so that $(L + n)v_n = 0$. Show from (d) above that $(A_1 + n)v_n$ is an eigenvector for L, with eigenvalue $-(n + 1)$. Conclude from (b) that if v_0 is an eigenvector of L with eigenvalue 0, then $A_n v_0 = v_n$ is an eigenvector with eigenvalue $-n$.

(iii) Show that if $P = d/dx$ and $Q =$ multiplication by x, then P and Q satisfy the relations above. Show also that

$$e^x \frac{d}{dx} e^{-x} = P - I.$$

Conclude that a solution to the Laguerre equation $zy'' + (1 - z)y' + ny = 0$ is

$$e^x \left(\frac{d}{dx}\right)^n (e^{-x} x^n) = (P - I)^n Q^n(1);$$

here 1 is the constant function on $\mathbb{R}$.

B.(i) Take P, Q and I as in A(i). Suppose $Pv_0 = 0$, and set $v_n = Q^n(v_0)$. Show inductively that $Pv_n = nv_{n-1}$. Conclude that v_n is an eigenvector of eigenvalue n for QP.

(ii) Put $P = d/dx, Q = (d/dx) + x$. Verify that these satisfy the correct commutation relations, and show that

$$Q = e^{-x^2/2} \frac{d}{dx} e^{x^2/2}.$$

Show that solutions of Hermite's equation $y'' + xy' - ny = 0$ are given by

$$y = e^{-x^2/2} \left(\frac{d}{dx}\right)^n e^{x^2/2} = Q^n(1).$$

In addition to Rodrigues-type formulas such as the above, one can deduce in a purely formal manner recursion relations and other properties of the Hermite, Laguerre, Legendre, Bessel, and many other classical families of functions.

I would like to thank Kenneth Gross for painstaking efforts to improve the readability of this paper.

References

[A] J. F. Adams, Lectures on Lie Groups, Benjamin, New York, 1969.

[Be] L. Bers, Riemann Surfaces, New York University Lecture Notes, 1957–58.

[Bo] A. Borel, Linear Algebraic Groups, Benjamin, New York, 1969.

[Bo2] A. Borel, On the development of Lie Group Theory, Math. Intelligencer, 2 (1980) 67–72.

[BC] A. Borel and W. Casselman, Editors, Automorphic Forms, Representations, and L-functions, Proc. Symp. Pure Math. XXXIII, American Mathematical Society, Providence, R.I., 1979.

[BT] A. Borel and J. Tits, Homomorphismes "abstraits" de groupes algébriques simples, Ann. of Math., (2) 97 (1973) 499–571.

[Br] M. Braun, Differential Equations and Their Applications, 2nd ed., Springer Verlag, New York, 1978.

[Ca] P. Cartier, Quantum mechanical commutation relations and theta functions, Proc. Symp. Pure Math. IX, American Mathematical Society, Providence, R.I., 1966, 361–383.

[Ch] C. Chevalley, Theory of Lie Groups I, Princeton University Press, Princeton, N.J., 1946.

[DN] J. Daniel, B. Noble, Applied Linear Algebra, 2nd ed., Prentice-Hall, Englewood Cliffs, N.J., 1977.

[Dy] F. J. Dyson, Mathematics in the Physical Sciences, Scientific American, 211(Sept. 1969) 128–147.

[E] A. Einstein, Zur Elektrodynamik bewegter Korper, Ann. Physik, Bd. 17 (1905) 891.

[F] R. Fenn, What is the geometry of a surface?, this MONTHLY, 90 (1983) 87–98.

[G] A. Gleason, Groups without small subgroups, Ann. of Math., 56 (1952) 193–212.

[Go] R. Godement, Introduction à la théorie des groupes de Lie (in 2 volumes), Pub. Math. de l'Univ. Paris VII, vol. 11, vol. 12, Paris, 1982.

[Gr] K. Gross, On the evolution of noncommutative harmonic analysis, this MONTHLY, 85 (1978) 525–528.

[Hg] W. Heisenberg, Uber quantentheoretische Umdeutung kinematischer und mechanischen Beziehungen, Phys., 33 (1925) 879.

[He] S. Helgason, Differential Geometry, Lie Groups, and Symmetric Spaces, Academic Press, New York, 1978.

[HP] E. Hille and R. Phillips, Functional Analysis and Semigroups, A.M.S. Colloquium Publications, vol. 31, American Mathematical Society, Providence, R.I., 1957.

[HS] M. Hirsch and S. Smale, Differential Equations, Dynamical Systems, and Linear Algebra, Academic Press, New York, 1974.

[Ho] R. Howe, On the role of the Heisenberg group in harmonic analysis, B.A.M.S. (New Series), 3 (1980) 821–843.

[Hu] J. Humphreys, Introduction to Lie algebras and representation theory, Springer Verlag, New York, 1972.

[Hn] T. Husain, Introduction to Topological Groups, Saunders, PA, 1966.

[Hr] D. Husemoller, Fibre Bundles, McGraw-Hill, New York, 1966.

[J1] N. Jacobson, Lie Algebras, Wiley-Interscience, New York, 1962.

[J2] N. Jacobson, Basic Algebra II, Freeman, San Francisco, 1980.

[Ka] I. Kaplansky, Lie Algebras and Locally Compact Groups, Chicago University Press, Chicago, 1971.

[Ke] J. Kelley, General Topology, Van Nostrand, Princeton, N.J., 1955.

[Kl] F. Klein, Vergleichende Betractungen uber neuere geometrische Forschungen, in Mathematische Abhandlungen, 1st ed., Springer-Verlag, Berlin, 1921.

[Ko] S. Kobayashi, Transformation Groups in Differential Geometry, Ergeb. Math., b. 70, Springer-Verlag, New York, Heidelberg, 1970.

[L] S. Lang, Analysis I, Addison-Wesley, Reading, MA, 1968.

[Lo] E. Loebl, Editor, Group Theory and Its Applications I, II, Academic Press, New York, 1968, 1971.

[Ma] G. Margulis, On the arithmeticity of nonuniform lattices, Proc. of Int. Cong. Math., Vancouver, 1970, pp. 21–35. (Russian).

[Ms] W. Massey, Algebraic Topology: An Introduction, Graduate Texts in Math., 56, Springer-Verlag, New York, Heidelberg, 1967.

[Me] A. Messiah, Quantum Mechanics II, North-Holland, Amsterdam, 1963.

[Mi] W. Miller, Lie theory and Special Functions, Academic Press, New York, 1968.

[Mk] H. Minkowski, Die Grundgleichungen fur die elektromagnetischen Vorgange in Bewegten Korpern, Nachr. d. K. Ges. d. Wissensch. zu Gottingen, 1908, 53.

[MZ] D. Montgomery and L. Zippin, Topological Transformation Groups, Wiley Interscience, New York, 1955.

[Mo] G. D. Mostow, Strong rigidity of locally symmetric spaces, Ann. of Math. Studies, no. 78, Princeton University Press, Princeton, N.J., 1973.

[N] M. Naimark, Normed Algebras, P. Noordhoff, N. V., Groningen, Netherlands, 1959.

[Nn] J. von Neumann, Uber die analytischen Eigenschaften Gruppen linearer Transformationen und ihrer Darstellungen, Math. Z., 30 (1929) 3–42.

[P] L. S. Pontryagin, Topological Groups, 2nd ed., Gordon and Breach, New York, 1966.

[Pr] R. Proctor, Solution of Two Difficult Combinatorial Problems with Linear Algebra, this MONTHLY, 89 (1982) 721–734.

[R] W. Rudin, Principles of Mathematical Analysis, 3rd ed., McGraw-Hill, New York, 1976.

[Se] J. P. Serre, Lie Algebras and Lie Groups, Benjamin, New York, 1965.

[St] G. Strang, Linear Algebra and its Applications, 2nd ed., Academic Press, New York, 1980.

[Th] W. Thurston, The Topology and Geometry of 3 Manifolds, Princeton University lecture notes.

[V] N. J. Vilenkin, Special Functions and the theory of group representations, A.M.S. Translations of Math. Monographs, no. 22, American Mathematical Society, Providence, R.I., 1968.

[Wg] E. Wigner, Symmetries and Reflections, Indiana University Press, Bloomington, IN, 1967.

[W] H. Weyl, Gruppentheorie und Quantenmechanik, S. Hirzel Verlag, Leipzig, 1928. (English translation: The Theory of Groups and Quantum Mechanics, Dover Publications, New York)