Research Article   **Open Access**

Matrix Lie Groups: An Introduction

Lawson J*

Boyd Professor, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803, USA

Abstract

This article presents basic notions of Lie theory in the context of matrix groups with goals of minimizing the required mathematical background and maximizing accessibility. It is structured with exercises that enhance the text and make the notes suitable for (part of) an introductory course at the upper level undergraduate or early graduate level. Indeed the notes were originally written as part of an introductory course to geometric control theory.

Introduction

Lie theory, the theory of Lie groups, Lie algebras, and their applications is a fundamental part of mathematics that touches on a broad spectrum of mathematics, including geometry (classical, differential, and algebraic), ordinary and partial differential equations, group, ring, and algebra theory, complex and harmonic analysis, number theory, and physics (classical, quantum, and relativistic). It typically relies upon an array of substantial tools such as topology, differentiable manifolds and differential geometry, covering spaces, advanced linear algebra, measure theory, and group theory to name a few. However, we will considerably simplify the approach to Lie theory by restricting our attention to the most important class of examples, namely those Lie groups that can be concretely realized as (multiplicative) groups of matrices.

Lie theory began in the late nineteenth century, primarily through the work of the Norwegian mathematician Sophus Lie, who called them "continuous groups," in contrast to the usually finite permutation groups that had been principally studied up to that point. An early major success of the theory was to provide a viewpoint for a systematic understanding of the newer geometries such as hyperbolic, elliptic, and projective, that had arisen earlier in the century. This led Felix Klein in his Erlanger Programme to propose that geometry should be understood as the study of quantities or properties left invariant under an appropriate group of geometric transformations. In the early twentieth century Lie theory was widely incorporated into modern physics, beginning with Einstein's introduction of the Lorentz transformations as a basic feature of special relativity. Since these early beginnings research in Lie theory has burgeoned and now spans a vast literature.

The essential feature of Lie theory is that one may associate with any Lie group G a Lie algebra $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is a vector space equipped with a bilinear non-associative anti-commutative product, called the *Lie bracket or commutator* and usually denoted $[\cdot,\cdot]$. The crucial and rather surprising fact is that a Lie group is almost completely determined by its Lie algebra $\mathfrak{g}$. There is also a basic bridge between the two structures given by the exponential map exp : $\mathfrak{g} \to G$. For many purposes structure questions or problems concerning the highly complicated nonlinear structure G can be translated and reformulated via the exponential map in the Lie algebra $\mathfrak{g}$, where they often lend themselves to study via the tools of linear algebra (in short, nonlinear problems can often be linearized). This procedure is a major source of the power of Lie theory.

The General Linear Group

Let V be a finite dimensional vector space equipped with a complete norm $|| \cdot ||$ over the field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. (Actually since the space V is finite dimensional, the norm must be equivalent to the usual euclidean norm, and hence complete.) Let End(V) denote the algebra of linear self-maps on V, and let GL(V) denote the *general linear group*, the group (under composition) of invertible self-maps. If $V = \mathbb{R}^n$, then End(V) may be identified with $M_n(\mathbb{R})$, the $n \times n$ matrices, and GL(V) = $GL_n(\mathbb{R})$, the matrices of nonvanishing determinant.

We endow End(V) with the usual *operator norm*, a complete norm defined by

$$\| A \| = \sup\{\| Av \| : \| v \| = 1\} = \sup\{\frac{\| Av \|}{\| v \|} : v \neq 0\},$$

which gives rise to the metric $d(A,B) = || B - A ||$ on End(V) and, by restriction, on GL(V).

Exercise 2.1: $\| AB \| \leq \| A \| \| B \|, \| tA \| = | t | \| A \|,$ and $\| A^n \| \leq \| A \|^n$

Exercise 2.2: Show that GL(V) is a dense open subset of End(V). (Hint: The determinant function is polynomial, hence continuous, and $A - (1/n)I$ converges to A and is singular for at most finitely many values, since the spectrum of A is finite.)

Exercise 2.3: The multiplication and inversion on GL(V) are analytic, i.e., expressible locally by power series. (Hint: the multiplication is actually polynomial and the cofactor expansion shows that inversion is rational.)

A group G endowed with a Hausdorff topology is called a *topological group* if the multiplication map $m : G \times G \to G$ and the inversion map on G are continuous. By the preceding exercise GL(V) is a topological group.

The Exponential Map

We define the *exponential map* on End V by

*Corresponding author: Lawson J, Boyd Professor, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803, United States, Tel: 12255781672; E-mail: lawson@math.lsu.edu

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

Lemma 1: *The exponential map is absolutely convergent, hence convergent on all of* End(V). *Hence it defines an analytic self-map on* End(V).

Proof: $\| \sum_{n=0}^{\infty} \frac{A^n}{n!} \| \leq \sum_{n=0}^{\infty} \frac{\| A \|^n}{n!} = \exp(\| A \|)$.

Absolute convergence allows us to rearrange terms and to carry out various algebraic operations and the process of differentiation term wise. We henceforth allow ourselves the freedom to carry out such manipulations without the tedium of a rather standard detailed verification.

Exercise 3.1: (i) Show that the exponential image of a block diagonal matrix with diagonal blocks $A_1,...,A_m$ is a block diagonal matrix with diagonal blocks $\exp(A_1),...,\exp(A_n)$. In particular, to compute the exponential image of a diagonal matrix, simply apply the usual exponential map to the diagonal elements.

(ii) Suppose that A is similar to a diagonal matrix, $A = PDP^{-1}$. Show that $\exp(A) = P \exp(D) P^{-1}$.

Proposition 2: If $A, B \in$ End V and $AB = BA$, then $\exp(A + B) = \exp A \exp B = \exp B \exp A$.

Proof: Computing term wise and rearranging we have

$$\exp A \exp B = \left(\sum_{n=0}^{\infty} \frac{A^n}{n!} \right) \left(\sum_{m=0}^{\infty} \frac{B^m}{m!} \right) = \sum_{n,m=0}^{\infty} \frac{A^n B^m}{n! m!}$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{n+m=k} \frac{k!}{n! m!} A^n B^m \right)$$

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

Since A and B commute, the familiar binomial theorem yields

$$(A + B)^k = \sum_{j=1}^{k} \binom{k}{j} A^j B^{k-j},$$

and substituting into the previous expression yields the proposition.

Let V, W be finite dimensional normed vector spaces and let $f: U \to W$, where U is a nonempty open subset of V. A linear map $L: V \to W$ is called the (*Frechet*) *derivative* of f at $x \in U$ if in some neighbourhood of x

$$f(x + h) - f(x) = L(h) + r(h), \text{ where } \lim_{h \to 0} \frac{r(h)}{\| h \|} = 0.$$

If it exists, the derivative is unique and denoted by $df(x)$ or $f'(x)$.

Lemma 3: *The identity map on* End(V) *is the derivative at 0 of* exp : End(V) $\to$ End(V), *i.e.,* $d \exp(0) = Id$.

Proof: For h $\in$ End(V), we have

$$\exp(h) - \exp(0) = \sum_{n=0}^{\infty} \frac{h^n}{n!} - 1_V = \sum_{n=1}^{\infty} \frac{h^n}{n!} = \text{Id}(h) + h^2 \sum_{n=2}^{\infty} \frac{h^{n-2}}{n!},$$

where $\lim_{h \to 0} \| \frac{h^2 \sum_{n=0}^{\infty} \frac{h^n}{(n+2)!}}{\| h \|} \| \leq \lim_{h \to 0} \frac{\| h \|^2}{\| h \|} \left(\sum_{n=2}^{\infty} \frac{1}{n!} \right) = 0.$

Applying the Inverse Function Theorem, we have immediately

from the preceding lemma

Proposition 4: *There exist neighbourhoods U of 0 and V of I in* End V *such that* $\exp |_U$ *is a diffeomorphism onto V.*

For $A \in$ End V and $r > 0$, let $B_r(A) = \{C \in V : \| C - A \| < r\}$.

Exercise 3.2: Show that $\exp(B_r(0)) \subseteq B_s(1_V)$ where $s = e^r - 1$. In particular for $r = \ln 2$, $\exp(B_r(0)) \subseteq B_1(1_V)$.

One-parameter Groups

A *one-parameter subgroup* of a topological group G is a continuous homomorphism $\alpha: \mathbb{R} \to G$ from the additive group of real numbers into G.

Proposition 5: *For V a finite dimensional normed vector space and* $A \in$ End V, *the map* $t \mapsto \exp(tA)$ *is a one-parameter subgroup of* GL(V). *In particular,* $(\exp(A))^{-1} = \exp(-A)$.

Proof: Since sA and tA commute for any $s, t \in \mathbb{R}$, we have from Proposition 2 that $t \mapsto \exp(tA)$ is a homomorphism from the additive reals to End V under multiplication. It is continuous, indeed analytic, since scalar multiplication and exp are. The last assertion follows from the homomorphism property and assures the image lies in GL(V).

Proposition 6: *Choose an* $r < \ln 2$. *Let* $A \in B_r(0)$ *and let* $Q = \exp A$. *Then* $P = \exp(A/2)$ *is the unique square root of Q contained in* $B_1(1_V)$.

Proof: Since $\exp(tA)$ defines a one-parameter subgroup,

$$P^2 = (\exp(A/2))^2 = \exp(A/2) \exp(A/2) = \exp(A/2 + A/2) = \exp(A) = Q.$$

Also $A \in B_r(0)$ implies $A/2 \in B_r(0)$, which implies $\exp(A/2) \in B_1(1_V)$ (Exercise 3.2).

Suppose two elements in $B_1(1_V)$, say $1+B$ and $1+C$ where $\| B \|, \| C \| < 1$ satisfy $(1 + B)^2 = (1 + C)^2$. Then expanding the squares, cancelling the 1's, and rearranging gives

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

Taking norms yields

$2 \| B - C \| \leq \| C \| \| C - B \| + \| C - B \| \| B \| = (\| C \| + \| B \|) \| C - B \|.$

This implies either $||C|| + ||B|| \geq 2$, which is also false since each summand is less than 1, or $||B - C|| = 0$, i.e., $B = C$. We conclude there at most one square root in $B_1(1_V)$.

Lemma 7: *Consider the additive group* $(\mathbb{R}, +)$ *of real numbers.*

(i) *If a subgroup contains a sequence of nonzero numbers* $\{a_n\}$ *converging to 0, then the subgroup is dense.*

(ii) *For one-parameter subgroups* $\alpha, \beta : \mathbb{R} \to G$, *the set* $\{t \in \mathbb{R} : \alpha(t) = \beta(t)\}$ *is a closed subgroup.*

Proof: (i) Let $t \in \mathbb{R}$ and let $\varepsilon > 0$. Pick a_n such that $|a_n| < \varepsilon$. Pick an integer k such that $|t/a_n - k| < 1$ (for example, pick k to be the floor of t/a_n). Then multiplying by $|a_n|$ yields $|t - ka_n| < |a_n| < \varepsilon$. Since ka_n must be in the subgroup, its density follows.

(ii) Exercise.

Exercise 4.1: Show that any nonzero subgroup of $(\mathbb{R}, +)$ is either dense or cyclic. (Hint: Let H be a subgroup and let $r = \inf\{t \in H : t > 0\}$. Consider the two cases $r = 0$ and $r > 0$.)

The next theorem is a converse of Proposition 5.

Theorem 8: *Every one parameter subgroup* $\alpha : \mathbb{R} \to$ End(V) *is of the form* $\alpha(t) = \exp(tA)$ *form some* $A \in$ End V.

Proof: Pick $r < ln\ 2$ such that exp restricted to $B_r(0)$ is a diffeomorphism onto an open subset containing $1 = 1_v$. This is possible by Proposition 4. Note that $\exp(B_r(0)) \subseteq B_1(1)$. By continuity of α, pick $0 < \varepsilon$ such that $\alpha(t) \in \exp(B_r(0))$ for all $-\varepsilon < t < \varepsilon$. Then $\alpha(1/2^k) \in \exp(B_r(0)) \subseteq B_1(1)$ for all $1/2^k < \varepsilon$.

Pick $1/2^n < \varepsilon$. Then $Q := \alpha(1/2^n) \in \exp(B_r(0))$ implies $Q = \alpha(1/2^n) = \exp(B)$ for some $B \in B_r(0)$. Set $A = 2^n B$. Then $Q = \exp((1/2^n)A)$.

Then $\alpha(1/2^{n+1})$ and $\exp(B/2)$ are both square roots of Q contained in $B_1(1)$, and hence by Proposition 6 are equal. Thus $\alpha(1/2^{n+1}) = \exp((1/2^{n+1})A)$. By induction $\alpha(1/2^{n+k}) = \exp((1/2^{n+k})A)$ for all positive integers k. By Lemma 7(ii) the two one-parameter subgroups agree on a closed subgroup, and by Lemma 7 this subgroup is also dense. Hence $\alpha(t)$ and $\exp(tA)$ agree everywhere.

The preceding theorem establishes that a merely continuous one-parameter subgroup must be analytic. This is a very special case of Hilbert's fifth problem, which asked whether a locally euclidean topological group was actually an analytic manifold with an analytic multiplication. This problem was solved positively some fifty years later in the 1950's by Gleason, Montgomery, and Zippin.

Exercise 4.2: Show that if $\exp(tA) = \exp(tB)$ for all $t \in \mathbb{R}$, then $A = B$. (Hint: Use Proposition 4)

Remark 9: The element $A \in \mathrm{End}\,V$ is called the *infinitesimal generator* of the one-parameter group $t \mapsto \exp(tA)$. We conclude from the preceding theorem and exercise that there is a one-to-one correspondence between one-parameter subgroups and their infinitesimal generators.

Curves in End V

In this section we consider basic properties of differentiable curves in End V. Let I be an open interval and let $A(\cdot) : I \to \mathrm{End}\ V$ be a curve. We say that A is C^r if each of the coordinate functions $A_{ij}(t)$ is C^r on $\mathbb{R}$. We define the derivative $\dot{A}(t) = \lim_{h \to 0}(1/h)(A(t+h) - A(t))$. The derivative exists iff the derivative $\dot{A}_{ij}(t)$ of each coordinate function exists, and in this case $\dot{A}(t)$ is the linear operator with coordinate functions $\frac{d}{dt}(A_{ij}(t))$.

Items (1) and (2) in the following list of basic properties for operator valued functions are immediate consequences of the preceding characterization, and item (5) is a special case of the general chain rule.

(1) $D_t(A(t) \pm B(t)) = \dot{A}(t) \pm \dot{B}(t)$

(2) $D_t(rA(t)) = r\dot{A}(t)$.

(3) $D_t(A(t) \cdot B(t)) = \dot{A}(t) \cdot B(t) + A(t) \cdot \dot{B}(t)$

(Note: Order is important since multiplication is noncommutative.)

(4) $D_t(A^{-1}(t)) = -A^{-1}(t) \cdot \dot{A}(t) \cdot A^{-1}(t)$.

(5) If $\dot{B}(t) = A(t)$, then $D_t(B(f(t))) = f'(t)A(f(t))$.

Exercise 5.1: Establish properties (3) and (4). (Hints: (3) Mimic the proof of the product rule in the real case. (4) Note $A^{-1}(t)$ is differentiable if $A(t)$ is, since it is the composition with the inversion function, which is analytic, hence C^r for all r. Differentiate the equation $A(t).A^{-1}(t) = I$ and solve for $D_t(A^{-1}(t))$.)

We can also define the integral $\int_a^b A(t)dt$ by taking the coordinate integrals $\int_a^b A_{ij}(t)dt$. The following are basic properties of the integral that follow from the real case by working coordinate wise.

(6) If $B(t) = \int_{t_0}^t A(s)ds$, then $\dot{B}(t) = A(t)$.

(7) If $\dot{B}(t) = A(t)$, then $\int_r^s A(t)dt = B(s) - B(r)$.

We consider curves given by power series: $A(t) = \sum_{n=0}^{\infty} t^n A_n$. Define the nth-partial sum to be $S_n(t) = \sum_{k=0}^n t^k A_k$. The power series *converges* for some value of t if the partial sums $S_n(t)$ converge in each coordinate to some $S(t)$. This happens iff the coordinatewise real power series all converge to the coordinates of $S(t)$.

Since for an operator A, $|a_{ij}| \leq \|A\|$ for each entry a_{ij} (exercise), we have that *absolute convergence*, the convergence of $\sum_{n=1}^{\infty} |t|^n \|A_n\|$, implies the absolute convergence of each of the coordinate series, and their uniform convergence over any closed interval in the open interval convergence of the real power series $\sum_{n=1}^{\infty} t^n \|A_n\|$. These observations justify term wise differentiation and integration in the interval of convergence of $\sum_{n=1}^{\infty} t^n \|A_n\|$.

Exercise 5.2: (i) Show that the power series

$$\exp(tA) = \sum_{n=0}^{\infty} \frac{t^n}{n!}A^n$$

is absolutely convergent for all t (note that $A_n = (1/n!)A^n$ in this series).

(ii) Use termwise differentiation to show $D_t(\exp(tA)) = A\exp(tA)$.

(iii) Show that $X(t) = \exp(tA)X_0$ satisfies the differential equation on End V given by

$$\dot{X}(t) = AX(t), \qquad X(0) = X_0.$$

(iv) Show that the equation $\dot{x}(t) = Ax(t), x(0) = x_0$ on V has solution $x(t) = \exp(tA)x_0$.

The Baker-Campbell-Hausdorff Formalism

It is a useful fact that the derivative of the multiplication map at the identity I of End V is the addition map.

Proposition 10: *Let* $m : \mathrm{End}(V) \times \mathrm{End}(V) \to \mathrm{End}(V)$ *be the multiplication map,* $m(A,B) = AB$. *Then the derivative at* (I,I), $dm_{(I,I)}$: $\mathrm{End}(V) \times \mathrm{End}(V) \to \mathrm{End}(V)$ *is given by* $dm_{(I,I)}(U, V) = U + V$.

Proof: Since the multiplication map is polynomial, continuous partials of all orders exist, and in particular the multiplication map is differentiable. By definition the value of the derivative at (I,I) evaluated at some $(U,0) \in \mathrm{End}(V) \times \mathrm{End}(V)$ is given by

$$dm_{(I,I)}(U,0) = \lim_{t \to 0} \frac{m(I+tU,I) - m(I,I)}{t} = \lim_{t \to 0} \frac{tU}{t} = U.$$

We have seen previously that the exponential function is a diffeomorphism from some open neighbourhood B of 0 to some open neighbourhood U of I. Thus there exists an analytic inverse to the exponential map, which we denote by $\log : U \to B$. Indeed if one defines

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

then just as for real numbers this series converges absolutely for $\|A\| < 1$. Further since $\exp(\log A) = A$ holds in the case of real numbers, it holds in the algebra of formal power series, and hence in the linear operator or matrix case. Indeed one can conclude that exp is 1 - 1 on $B_{\ln 2}(0)$, carries it into $B_1(I)$, and has inverse given by the preceding logarithmic series, all this without appeal to the Inverse Function Theorem.

The local diffeomorphic property of the exponential function allows one to pull back the multiplication in GL(V) locally to a neighbourhood of 0 in End V. One chooses two points A, B in a sufficiently small neighbourhood of 0, forms the product $\exp(A) \cdot \exp(B)$ and takes the log of this product:

$$A * B := \log(\exp A \cdot \exp B)$$

This *Baker-Campbell-Hausdorff multiplication* is defined on any $B_r(0)$ small enough so that $\exp(B_r(0)) \cdot \exp(B_r(0))$ is contained in the domain of the log function; such exist by the local diffeomorphism property and the continuity of multiplication. Now there is a beautiful formula called the *Baker-Campbell-Hausdorff formula* that gives $A * B$ as a power series in A and B with the higher powers being given by higher order Lie brackets or commutators, where the (first-order) *commutator* or *Lie bracket* is given by $[A,B] := AB - BA$. The Baker-Campbell-Hausdorff power series is obtained by manipulating the power series for $\log(\exp(x) \cdot \exp(y))$ in two noncommuting variables x, y in such a way that it is rewritten so that all powers are commutators of some order. To develop this whole formula would take us too far afield from our goals, but we do derive the first and second order terms, which suffice for many purposes.

Definition 11: An open ball $B_r(0)$ is called a *Baker-Campbell-Hausdorff neighbourhood*, or BCH-*neighborhood* for short, if $r < 1/2$ and $\exp(B_r(0) \cdot \exp(B_r(0)) \subseteq B_s(0)$ for some s,r such that exp restricted to $B_s(0)$ is a diffeomorphism onto some open neighbourhood of I. By the local diffeomorphism property of the exponential map and the continuity of multiplication at I, BCH-neighbourhoods always exist. We define the *Baker-Campbell-Hausdorff multiplication* on any BCH-neighbourhood $B_r(0)$ by

$$A * B = \log(\exp A \cdot \exp B).$$

Note that $A * B$ exists for all $A, B \in B_r(0)$, but we can only say that $A * B \in \text{End } V$, not necessarily in $B_r(0)$.

Proposition 12: Let $B_r(0)$ be a *BCH-neighbourhood*. Define $\mu : B_r(0) \times B_r(0) \to \text{End } V$ by $\mu(A,B) = A * B$. Then

(i) $A * B = A + B + R(A,B)$ where $\displaystyle\lim_{A,B\to 0} \frac{\|R(A,B)\|}{\|A\| + \|B\|} = 0$.

(ii) There exists $0 < s \le r$ such that $\|A*B\| \le 2(\|A\| + \|B\|)$ for $A,B \in B_s(0)$.

Proof: (i) We have that $\mu = \log \circ m \circ (\exp \times \exp)$ so by the chain rule, the fact that the derivatives of exp at 0 and log at I are both the identity map $Id : \text{End } V \to \text{End } V$ (Lemma 3 and the Inverse Function Theorem) and Proposition 10, we conclude that

$$d\mu_{(0,0)}(U,V) = Id \circ dm_{(I,I)} \circ (Id \times Id)(U,V) = U + V.$$

By definition of the derivative, we have

(1) $U * V = U * V - 0 * 0 = dm_{(0,0)}(U,V) + R(U,V)$

$\qquad = U + V + R(U,V)$ where $\displaystyle\lim_{(U,V)\to(0,0)} \frac{\|R(U,V)\|}{\|U\| + \|V\|} = 0$.

(Note that the second equality is just the definition of the derivative, where the norm on End $V \times$ End V is the sum norm.) This gives (i).

(ii) Using (i), we obtain the following string:

$$\|A * B\| \le \|A * B - A - B\| + \|A + B\| \le \|R(A,B)\| + \|A\| + \|B\|.$$

Now $\|R(A,B)\| \to 0$ as $A, B \to 0$, so the right-hand sum is less than or equal $2(\|A\| + \|B\|)$ on some $B_s(0) \subseteq B_r(0)$.

Exercise 6.1: Use the fact that $0 * 0 = 0$ and imitate the proof of Proposition 10 to show directly that $dm_{(0,0)}(U,V) = U + V$.

We now derive the linear and second order terms of the Baker-Campbell-Hausdorff series.

Theorem 13: Let $B_r(0)$ be a BCH-neighbourhood. Then

$$A * B = A + B + \frac{1}{2}[A,B] + S(A,B) \text{ where } \lim_{A,B\to 0} \frac{\|S(A,B)\|}{(\|A\| + \|B\|)^2} = 0.$$

Proof: Pick $B_s(0) \subseteq B_r(0)$ so that condition (ii) of Proposition 12 is satisfied. Setting $C = A * B$, we have directly from the definition of $A * B$ that $\exp C = \exp A \cdot \exp B$. By definition

(2) $\exp C = I + C + \dfrac{C^2}{2} + R(C)$, where $R(C) = \displaystyle\sum_{n=3}^{\infty} \frac{C^n}{n!}$.

For $A, B \in B_s(0)$, we have from Proposition 12 that $\|C\| = \|A * B\| \le 2(\|A\| + \|B\|) < 1$ since r < 1/2. Thus we have the estimate

(3) $\|R(C)\| \le \displaystyle\sum_{n=3}^{\infty} \frac{\|C\|^n}{n!} \le \|C\|^3 \sum_{n=3}^{\infty} \frac{\|C\|^{n-3}}{n!} \le \frac{1}{2}\|C\|^3$.

Recalling the calculations in the proof of Proposition 2, we have

(4) $\exp A \cdot \exp B = I + A + B + \dfrac{1}{2}(A^2 + 2AB + B^2) + R_2(A,B)$,

where $R_2(A,B) = \displaystyle\sum_{n=3}^{\infty} \frac{1}{n!}\Big(\sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}\Big)$.

We have the estimate

(5) $\|R_2(A,B)\| \le \displaystyle\sum_{n=3}^{\infty} \frac{1}{n!}(\|A\| + \|B\|)^n \le \frac{1}{2}(\|A\| + \|B\|)^3$.

If in the equation $\exp C = \exp A \cdot \exp B$, we replace $\exp C$ by the right side of equation (2), $\exp A \cdot \exp B$ by the right side equation (4), and solve for C, we obtain

(6) $C = A + B + \dfrac{1}{2}(A^2 + 2AB + B^2 - C^2) + R_2(A,B) - R(C)$.

Since

$A^2 + 2AB + B^2 - C^2 = [A,B] + (A + B)^2 - C^2$

$\qquad\qquad\qquad\qquad = [A,B] + (A + B)(A + B - C) + (A + B - C)C,$

we obtain alternatively

(7) $C = A + B + \dfrac{1}{2}[A,B] + S(A,B)$,

where $S(A,B) = \dfrac{1}{2}((A + B)(A + B - C) + (A + B - C)C) + R_2(A,B) - R(C)$.

To complete the proof, it suffices to show that the limit as $A, B \to 0$ of each of the terms of $S(A,B)$ divided by $(\|A\| + \|B\|)^2$ is 0. First we have in $B_s(0)$

$$\frac{1}{2}\|(A+B)(A+B-C) + (A+B-C)C\| \leq \frac{1}{2}(\|A\| + \|B\| + \|C\|)\|A+B-C\|$$

$$\leq \frac{1}{2}\|A\| + \|B\| + 2(\|A\| + \|B\|))\|R(A,B)\|$$

$$= \frac{3}{2}(\|A\| + \|B\|)\|R(A,B)\|,$$

where the second inequality and last equality follow by applying appropriate parts of Proposition 12. Proposition 12 also insures that

$$\lim_{A,B \to 0} \frac{3(\|A\| + \|B\|)\|R(A,B)\|}{2(\|A\| + \|B\|)^2} = 0.$$

That $\lim_{A,B \to 0} \|R_2(A,B)\|/(\|A\| + \|B\|)^2 = 0$ follows directly from equation (5). Finally by equation (3) and Proposition 12(ii)

$$\frac{\|R(C)\|}{(\|A\| + \|B\|)^2} \leq \frac{1}{2}\frac{\|C\|^3}{(\|A\| + \|B\|)^2} \leq \frac{4(\|A\| + \|B\|)^3}{(\|A\| + \|B\|)^2}$$

which goes to 0 as $A, B \to 0$.

The Trotter and Commutator Formulas

In the following sections we show that one can associate with each closed subgroup of $GL(V)$ a Lie subalgebra of End V, that is, a subspace closed under Lie bracket. The exponential map carries this Lie algebra into the matrix group and using properties of the exponential map, one can frequently transfer structural questions about the Lie group to the Lie algebra, where they often can be treated using methods of linear algebra. In this section we look at some of the basic properties of the exponential map that give rise to these strong connections between a matrix group and its Lie algebra.

Theorem 14: (*Trotter Product Formula*) *Let $A,B \in$ End V and let $lim_n nA_n = A, lim_n nB_n = B$. Then*

(i) $A + B = \lim_n n(A_n * B_n)$;

(ii) $\exp(A + B) = \lim_n (\exp A_n \exp B_n)^n = \lim_n (\exp(A/n)\exp(B/n))^n$.

Proof: (i) Let $\varepsilon > 0$. For large n, $n\|A_n\| \leq \|A\| + \|nA_n - A\| < \|A\| + \varepsilon$, and thus $\|A_n\| \leq (1/n)(\|A\| + \varepsilon)$. It follows that $\lim_n A_n = 0$ and similarly $\lim_n B_n = 0$. By Proposition 12(i) we have

$$\lim_n n(A_n * B_n) = \lim_n nA_n + \lim_n nB_n + \lim_n nR(A_n, B_n) = A + B,$$

provided that $\lim_n nR(A_n, B_n) = 0$. But we have

$$\|nR(A_n, B_n)\| = \frac{n(\|A_n\| + \|B_n\|)\|R(A_n, B_n)\|}{\|A_n\| + \|B_n\|} \to (\|A\| + \|B\|) \cdot 0 \text{ as } n \to \infty.$$

(ii) The first equality follows directly by applying the exponential function to (i):

$$\exp(A + B) = \exp(\lim_n n(A_n * B_n)) = \lim_n \exp(n(A_n * B_n))$$

$$= \lim_n (\exp(A_n * B_n))^n = \lim_n (\exp(A_n)\exp(B_n))^n$$

where the last equality follows from the fact that exp is a local isomorphism from the BCH-multiplication to operator multiplication, and penultimate equality from the fact that $\exp(nA) = \exp(A)^n$, since exp restricted to $\mathbb{R}A$ is a one-parameter group. The second equality in part (ii) of the theorem follows from the first by setting $A_n = A/n$, $B_n = B/n$.

The exponential image of the Lie bracket of the commutator can be calculated from products of group commutators.

Theorem 15: (*Commutator Formula*) *Let $A, B \in$ End V and let $\lim_n nA_n = A, \lim_n nB_n = B$. Then*

(i) $[A, B] = \lim_n n^2(A_n * B_n - B_n * A_n) = \lim_n n^2(A_n * B_n * (-A_n) * (-B_n))$;

(ii) $\exp[A,B] = \lim_n (\exp(A_n)\exp(B_n)(\exp A_n)^{-1}(\exp B_n)^{-1})^{n^2}$

$$= \lim_n (\exp(A/n)\exp(B/n)(\exp(A/n))^{-1}(\exp(B/n))^{-1})^{n^2}.$$

Proof. (i) From Theorem 13 we have for A, B in a BCH-neighbourhood:

$$A * B - B * A = \frac{1}{2}([A,B] - [B,A]) + S(A,B) - S(B,A)$$

$$= [A,B] + S(A,B) - S(B,A)$$

since $[A, B] = -[B, A]$. Therefore

$$\lim_n n^2(A_n * B_n - B_n * A_n) = \lim_n n^2([A_n, B_n] + S(A_n, B_n) - S(B_n, A_n))$$

$$= \lim_n [nA_n, nB_n] + \lim_n (n^2 S(A_n, B_n) - n^2 S(B_n, A_n)),$$

$$= [A, B],$$

provided $\lim_n n^2 S(A_n, B_n) = \lim_n n^2 S(B_n, A_n)) = 0$. To see this, we note

$$\lim_n n^2\|S(A_n, B_n)\| = \lim_n n^2(\|A_n\| + \|B_n\|)^2 \frac{\|S(A_n, B_n)\|}{(\|A_n\| + \|B_n\|)^2} = (\|A\| + \|B\|)^2 \cdot 0 = 0$$

and similarly $\lim_n n^2\|S(B_n, A_n)\| = 0$

To see second equality in item (i), observe first that on a BCH-neighbourhood where the exponential map is injective,

$$\exp((-A) * (-B)) = \exp(-A)\exp(-B) = (\exp A)^{-1}(\exp B)^{-1}$$

$$= ((\exp B)(\exp A))^{-1} = (\exp(B * A))^{-1} = \exp(-B * A),$$

which implies $(-A) * (-B) = -(B * A)$. Hence we have by Theorem 13 that

$$A * B * (-A) * (-B) - (A * B - B * A) = (A * B) * (-(B * A)) - (A * B + (-B * A))$$

$$= \frac{1}{2}[A * B, -B * A] + S(A * B, -B * A).$$

Applying this equality to the given sequences, we obtain

$$n^2\|A_n * B_n * (-A_n) * (-B_n) - (A_n * B_n - B_n * A_n)\|$$

$$\leq \frac{n^2}{2}\|[A_n * B_n, -B_n * A_n]\| + n^2\|S(A_n * B_n, -B_n * A_n)\|.$$

Now if we show that the two terms in the second expression approach 0 as $n \to \infty$, then the first expression approaches 0, and thus the two limits in (i) will be equal. We observe first that by the Trotter Product Formula

$$\lim_n [A_n * B_n, -B_n * A_n] = \lim_n [nA_n * B_n, -nB_n * A_n] = [A + B, -(B + A)] = 0$$

since $[C, -C] = -[C, C] = 0$ for any C. Thus the first right-hand term approaches 0. For the second

$$n^2\|S(A_n * B_n, -B_n * A_n)\|$$

$$= n^2(\|A_n * B_n\| + \|-B_n * A_n\|)^2 \frac{\|S(A_n * B_n, -B_n * A_n)\|}{(\|A_n * B_n\| + \|-B_n * A_n\|)^2}$$

$$\to (\|A + B\| + \|-(B + A)\|)^2 \cdot 0 = 0$$

as $n \to \infty$.

(ii) The proof follows from an application of the exponential function to part (i), along the lines of the Trotter Product Formula.

Exercise 7.1: Give the proof of part (ii) in the preceding theorem.

The Lie Algebra of a Matrix Group

In this section we set up the fundamental machinery of Lie theory, namely we show how to assign to each matrix group a (uniquely determined) Lie algebra and an exponential map from the Lie algebra to the matrix group that connects the two together. We begin by defining the notions and giving some examples.

By a *matrix group* we mean a closed subgroup of GL(V), where V is a finite dimensional vector space.

Examples 8.1: The following are standard and basic examples.

(1) The general linear group GL(V). If $V = \mathbb{R}^n$, then we write the group of $n \times n$ invertible matrices as $GL_n(\mathbb{R})$.

(2) The special linear group $\{g \in GL(V) : \det(g) = 1\}$.

(3) Let V be a real (resp. complex) Hilbert space equipped with an inner product $\langle \cdot , \cdot \rangle$. The orthogonal group (resp. unitary group) consists of all transformations preserving the inner product, i.e.,

$$O(V)(\text{ resp. } U(V)) = \{g \in GL(V) : \forall x, y \in V, \langle gx, gy \rangle = \langle x, y \rangle\}$$

If $V = \mathbb{R}^n$ (resp. $\mathbb{C}^n$) equipped with the usual inner product, then the orthogonal group O_n (resp. unitary group U_n) consists of all $g \in GL(V)$ such that $g^t = g^{-1}$ (resp. $g^* = g^{-1}$).

(4) Let $V = \mathbb{R}^n \oplus \mathbb{R}^n$ equipped with the sympletic form

$$Q\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) := \langle x_1, y_2 \rangle - \langle y_1, x_2 \rangle.$$

The *real sympletic group* is the subgroup of GL(V) preserving Q:

$$\text{Sp}(V) = SP_{2n}(\mathbb{R}) := \{M \in GL(V) : \forall x, y \in V, Q(Mx, My) = Q(x, y)\}.$$

(5) Let $0 < m, n$ and consider the group of block upper triangular real matrices

$$U_{m,n} = \left\{ \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \in GL_{m+n}(\mathbb{R}) : A \in GL_m(\mathbb{R}), B \in M_{m,n}(\mathbb{R}), D \in GL_n(\mathbb{R}) \right\}.$$

This is the subgroup of $GL_{m+n}(\mathbb{R})$ that carries the subspace $\mathbb{R}^m \oplus \{0\}$ of $\mathbb{R}^m \oplus \mathbb{R}^n$ into itself.

Exercise 8.1. (i) Verify that the subgroups in (2)-(5) are closed.

(ii) Verify the alternative characterizations of elements of the subgroup in items (3) and (5).

Exercise 8.2. Establish the following equivalence:

$M \in SP_{2n}(\mathbb{R})$;

$M^* JM = J$ where $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \in \text{End}(\mathbb{R}^{2n})$;

If M has block matrix form $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ (where all submatrices are $n \times n$), then

$A^* C, B^* D$ are symmetric, and $A^* D - C^* B = I$

Definition 16. A real Lie algebra $\mathfrak{g}$ is a real vector space equipped with a binary operation

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

satisfying the identities

(i) (Bilinearity) For all $\lambda, \mu \in \mathbb{R}$ and $X, Y, Z \in \mathfrak{g}$,

$$[\lambda X + \mu Y, Z] = \lambda[X, Z] + \mu[Y, Z]$$

$$[X, \lambda Y + \mu Z] = \lambda[X, Y] + \mu[X, Z].$$

(ii) (Skew symmetry) For all $X, Y \in \mathfrak{g}$

$$[X, Y] = -[Y, X];$$

(iii) (Jacobi identity) For all X, Y, Z $\in \mathfrak{g}$,

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Exercise 8.3. Verify that End V equipped with the Lie bracket or commutator operation $[A, B] = AB - BA$ is a Lie algebra.

It follows directly from the preceding exercise that any subspace of End V that is closed with respect to the Lie bracket operation is a Lie subalgebra.

We define a *matrix semigroup S* to be a closed multiplicative sub-semigroup of $GL(V)$ that contains the identity element. We define the *tangent set of S* by

$$L(S) = \{A \in \text{End} V : \exp(tA) \in S \text{ for all } t \geq 0\}.$$

We define a *wedge* in End V to be a closed subset containing $\{0\}$ that is closed under addition and scalar multiplication by nonnegative scalars.

Proposition 17. *If S is a matrix semigroup, then L(S) is a wedge.*

Proof. Since $I = \exp(t.0)$ for all $t \geq 0$ and $I \in S$, we conclude that $0 \in L(S)$. If $A \in L(S)$, then $\exp(tA) \in S$ for all $t \geq 0$, and thus $\exp(rtA) \in S$ for all $r, t \geq 0$ It follows that $rA \in L(S)$ for $r \geq 0$. Finally by the Trotter Product Formula if $A, B \in L(S)$, then

$$\exp(t(A + B)) = \lim_n(\exp(tA/n)\exp(tB/n))^n \in S \text{ for } t \geq 0$$

since S is a closed subsemigroup of GL(V). Thus $A + B \in L(S)$.

Theorem 18. *For a matrix group $G \subseteq$ GL(V), the set*

$$\mathfrak{g} = \{A \in \text{End} V : \exp(tA) \in G \text{ for all } t \in \mathbb{R}\}.$$

is a Lie algebra, called the Lie algebra of G.

Proof. As in the proof of Proposition 17, $\mathfrak{g}$ is closed under addition and scalar multiplication, i.e., a subspace of End V. By the Commutator Formula for $A, B \in \mathfrak{g}$,

$$\exp([A, B]) = \lim_n((\exp A/n)(\exp B/n)(\exp(-A/n)(\exp(-B/n))^{n^2} \in G$$

since G is a closed subgroup of GL(V). Replacing A by tA, which again is in $\mathfrak{g}$, we have $\exp(t[A, B]) = \exp([tA, B]) \in G$ for all $t \in \mathbb{R}$. Thus $[A, B] \in \mathfrak{g}$.

Exercise 8.4. Show for a matrix group G (which is a matrix semi-group, in particular) that $\mathfrak{g} = L(G)$.

Lemma 19. *Suppose that G is a matrix group, $\{A_n\}$ is a sequence in End V such that $A_n \to 0$ and $\exp(A_n) \in G$ for all n. If $s_n A_n$ has a cluster point for some sequence of real numbers s_n, then the cluster point belongs to $\mathfrak{g}$.*

Proof. Let B be a cluster point of $s_n A_n$. By passing to an appropriate subsequence, we may assume without loss of generality that

$s_n A_n$ converges to B. Let $t \in \mathbb{R}$ and for each n pick an integer m_n such that $| m_n - ts_n | < 1$. Then

$$\| m_n A_n - tB \| = \| (m_n - ts_n) A_n + t(s_n A_n - B) \|$$
$$\leq | m_n - ts_n | \| A_n \| + | t | \| s_n A_n - B \|$$
$$\leq \| A_n \| + | t | \| s_n A_n - B \| \to 0$$

which implies $m_n A_n \to tB$. Since $\exp(m_n A_n) = (\exp A_n)^{m_n} \in G$ for each n and G is closed, we conclude that the limit of this sequence $\exp(tB)$ is in G. Since t was arbitrary, we see that $B \in \mathfrak{g}$.

We come now to a crucial and central result.

Theorem 20. *Let $G \subseteq \mathrm{GL}(V)$ be a matrix group. Then all sufficiently small open neighborhoods of 0 in $\mathfrak{g}$ map homeomorphically onto open neighborhoods of I in G.*

Proof. Let $B_r(0)$ be a BCH-neighborhood around 0 in $\mathrm{End}\ V$, which maps homeomorphically under $\exp$ to an open neighborhood $\exp(B_r(0))$ of I in $\mathrm{GL}(V)$ with inverse $\log$. Assume that $\exp(B_r(0) \cap \mathfrak{g})$ does not contain a neighborhood of I in G. Then there exists a sequence g_n contained in G but missing $\exp(B_r(0) \cap \mathfrak{g})$ that converges to I. Since $\exp(B_r(0))$ is an open neighborhood of I, we may assume without loss of generality that the sequence is contained in this open set. Hence $A_n = \log g_n$ is defined for each n, and $A_n \to 0$. Note that $A_n \in B_r(0)$, but $A_n \notin \mathfrak{g}$, for each n, since otherwise $\exp(A_n) = g_n \in \exp(\mathfrak{g} \cap B_r(0))$.

Let W be a complementary subspace to $\mathfrak{g}$ in $\mathrm{End}\ V$ and consider the restriction of the BCH-multiplication $\mu(A, B) = A * B$ to $(\mathfrak{g} \cap B_r(0)) \times (W \cap B_r(0))$. By the proof of Proposition 12, the derivative $d\mu_{(0,0)}$ of μ at $(0, 0)$ is addition, and so the derivative of the restriction of μ to $(\mathfrak{g} \cap B_r(0)) \times (W \cap B_r(0))$ is the addition map $+ : \mathfrak{g} \times W \to \mathrm{End}\ V$. Since $\mathfrak{g}$ and W are complementary subspaces, this map is an isomorphism of vector spaces. Thus by the Inverse Function Theorem there exists an open ball $B_s(0)$, $0 < s \leq r$, such that μ restricted to $(\mathfrak{g} \cap B_s(0)) \times (W \cap B_s(0))$ is a diffeomorphism onto an open neighborhood Q of 0 in $\mathrm{End}\ V$. Since $A_n \in Q$ for large n, we have $A_n = B_n * C_n$ (uniquely) for $B_n \in (\mathfrak{g} \cap B_s(0))$ and $C_n \in (W \cap B_s(0))$. Since the restriction of μ is a homeomorphism and $0 * 0 = 0$, we have $(B_n, C_n) \to (0, 0)$, i.e., $B_n \to 0$ and $C_n \to 0$.

By compactness of the unit sphere in $\mathrm{End}\ V$, we have that $C_n/\|C_n\|$ clusters to some $C \in W$ with $\|C\| = 1$. Furthermore,

$$g_n = \exp(A_n) = \exp(B_n * C_n) = \exp(B_n) \exp(C_n)$$

so that $\exp(C_n) = (\exp B_n)^{-1} g_n \in G$. It follows from Lemma 19 that $C \in \mathfrak{g}$. But this is impossible since $\mathfrak{g} \cap W = \{0\}$ and $C \neq 0$. We conclude that $\exp(B_r(0) \cap \mathfrak{g})$ does contain some neighborhood N of I in G.

Pick any open neighborhood $U = (B_r(0) \cap \mathfrak{g})$ of 0 in $\mathfrak{g}$ such that $\exp(U) \subseteq N$. Then $\exp U$ is open in $\exp(B_r(0) \cap \mathfrak{g})$ (since $\exp$ restricted to $B_r(0)$ is a homeomorphism), hence is open in N, and thus is open in G, being an open subset of an open set.

We sketch here how that theory of matrix groups develops from what we have already done in that direction. Recall that a manifold is a topological space M, which we will assume to be metrizable, that has a covering of open sets each of which is homeomorphic to an open subset of euclidean space. Any family of such homeomorphisms from any open cover of M is called an atlas, and the members of the atlas are called charts. The preceding theorem allows us to introduce charts on a matrix group G in a very natural way. Let U be an open set around 0 in $\mathfrak{g}$ contained in a BCH-neighborhood such that $W = \exp U$ is an open neighborhood of I in G. Let $\lambda_g : G \to G$ be the left translation map,

i.e., $\lambda_g(h) = gh$. We define an atlas of charts on G by taking all open sets $g^{-1}N$, where N is an open subset of G such that $I \in N \subseteq W$ and defining the chart to be $\log \lambda_g : g^{-1}N \to \mathfrak{g}$ (to view these as euclidean charts, we identify $\mathfrak{g}$ with some $\mathbb{R}^n$ via identifying some basis of $\mathfrak{g}$ with the standard basis of $\mathbb{R}^n$). One can check directly using the fact that multiplication of matrices is polynomial that for two such charts and φ and ψ, the composition $\varphi \circ \psi^{-1}$, where defined, is smooth, indeed analytic. This gives rise to a differentiable structure on G, making it a smooth (analytic) manifold. The multiplication and inversion on G, when appropriately composed with charts are analytic functions, and thus one obtains an analytic group, a group on an analytic manifold with analytic group operations. This is the unique analytic structure on the group making it a smooth manifold so that the exponential map is also smooth.

The Lie Algebra Functor

We consider the category of matrix groups to be the category with objects matrix groups and morphisms continuous (group) homomorphisms and the category of Lie algebras with objects subalgebras of some $\mathrm{End}\ V$ and morphisms linear maps that preserve the Lie bracket., The next result shows that the assignment to a matrix group of its Lie algebra is functorial.

Proposition 21. *Let $\alpha : G \to H$ be a continuous homomorphism between matrix groups. Then there exists a unique Lie algebra homomorphism*

$d\alpha : \mathfrak{g} \to \mathfrak{h}$ *such that the following diagram commutes*:

$$
\begin{array}{ccc}
G & \xrightarrow{\ \alpha\ } & H \\
\exp \uparrow & & \uparrow \exp \\
\mathfrak{g} & \xrightarrow{\ d\alpha\ } & \mathfrak{h}.
\end{array}
$$

Proof. Let $A \in \mathfrak{g}$. Then the map $\beta(t) := \alpha(\exp(tA))$ is a one-parameter subgroup of H. Hence it has a unique infinitesimal generator $\tilde{A} \in \mathfrak{h}$ Define $d\alpha(A) = \tilde{A}$. We show that $d\alpha$ is a Lie algebra homomorphism.

For $r \in \mathbb{R}$,

$\alpha(\exp(trA)) = \exp(tr\tilde{A})$,

so the infinitesimal generator for the left-hand side is $r\tilde{A}$. This shows that $d\alpha(rA) = r\tilde{A} = rd\alpha(A)$, so $d\alpha$ is homogeneous.

Let $A, B \in G$. Then

$$(\alpha \circ \exp)(t(A + B)) = (\alpha \circ \exp)(tA + tB) = \alpha(\lim_n (\exp(tA/n)\exp(tB/n))^n)$$
$$= \lim_n (\alpha(\exp(tA/n))\alpha(\exp(tB/n)))^n$$
$$= \lim_n (\exp(t\tilde{A}/n)\exp(t\tilde{B}/n))^n$$
$$= \exp(t\tilde{A} + t\tilde{B}) = \exp(t(\tilde{A} + \tilde{B})).$$

This shows that $d\alpha(A + B) = \tilde{A} + \tilde{B} = d\alpha(A) + d\alpha(B)$, and thus $d\alpha$ is linear. In an analogous way using the commutator, one shows that $d\alpha$ preserves the commutator.

If $d\alpha(A) = \tilde{A}$, then by definition for all t, $\alpha(\exp(tA)) = \exp(t\tilde{A})$.

For $t = 1$, $\alpha(\exp A) = \exp(\tilde{A}) = \exp(d\alpha(A))$. Thus $\alpha \circ \exp = d\alpha \circ \exp$. This shows the square commutes. If $\gamma : \mathfrak{g} \to \mathfrak{h}$ is another Lie algebra homomorphism that also makes the square commute, then for $A \in \mathfrak{g}$ and all $t \in \mathbb{R}$,

$$\exp(t d\alpha(A)) = \exp(d\alpha(tA)) = \alpha(\exp(tA)) = \exp(\gamma(tA)) = \exp(t\gamma(A)).$$

The uniqueness of the infinitesimal generator implies $d\alpha(A) = \gamma(A)$, and hence $d\alpha = \gamma$.

Exercise 9.1. Show that $d\alpha$ preserves the commutator.

Exercise 9.2. Let $\alpha : G \to H$ be a continuous homomorphism of matrix groups. Then the kernel K of α is a matrix group with Lie algebra the kernel of $d\alpha$.

10. Computing Lie Algebras

In this section we consider some tools for computing the Lie algebra of a matrix group, or more generally a closed subsemigroup of a matrix group. We begin with a general technique.

Proposition 22. *Let $\beta(\cdot, \cdot)$ be a continuous bilinear form on V and set*

$$G = \{g \in \mathrm{GL}(V) : \forall x, y \in V, \beta(gx, gy) = \beta(x, y)\}.$$

Then

$$\mathfrak{g} = \{A \in \mathrm{End}\ V : \forall x, y \in V, \beta(Ax, y) + \beta(x, Ay) = 0\}.$$

Proof. If $A \in g$, then $\beta(\exp(tA)x, \exp(tA)y) = \beta(x, y)$ for all $x, y \in V$. Differentiating the equation with respect to t by the product rule (which always holds for continuous bilinear forms), we obtain

$$\beta(A\exp(tA)x, \exp(tA)y) + \beta(\exp(tA)x, A\exp(tA)y) = 0.$$

Evaluating at $t = 0$ yields $\beta(Ax, y) + \beta(x, Ay) = 0$.

Conversely suppose for all $x, y \in V$, $\beta(Ax, y) + \beta(x, Ay) = 0$. Then from the computation of the preceding paragraph the derivative of

$$f(t) := \beta(\exp(tA)x, \exp(tA)y)$$

is $f'(t) = 0$. Thus f is a constant function with the value $\beta(x, y)$ at 0. It follows that $\exp(tA) \in G$ for all t, i.e., $A \in g$.

Exercise 10.1. Apply the preceding proposition to show that the Lie algebra of the orthogonal group $O_n(\mathbb{R})$ (resp. the unitary group $U_n(\mathbb{C})$) is the Lie algebra of $n \times n$ skew symmetric (skew hermitian) matrices.

Exercise 10.2. (i) Use the Jordan decomposition to show for any $A \in M_n(\mathbb{C})$, $\exp(\mathrm{tr}\ A) = \det(\exp A)$.

(ii) Use (i) and Exercise 9.2 to show that the Lie algebra of the group $SL_n(\mathbb{C})$ of complex matrices of determinant one is the Lie algebra of matrices of trace 0. (Hint: the determinant mapping is a continuous homomorphism from $GL_n(\mathbb{C})$ to the multiplicative group of non-zero complex numbers.)

(iii) Observe that $L(G \cap H) = L(G) \cap L(H)$. What is the Lie algebra of $SU_n(\mathbb{C})$, the group of unitary matrices of determinant one?

Exercise 10.3. Let $V = \mathbb{R}^n \oplus \mathbb{R}^n$ equipped with the canonical sympletic form

$$Q\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) := \langle x_1, y_2 \rangle - \langle y_1, x_2 \rangle.$$

The Lie algebra of $\mathrm{Sp}(V)$ is given by

$$\mathfrak{sp}(V) = \left\{ \begin{bmatrix} A & B \\ C & D \end{bmatrix} : D = -A^*, B = B^*, C = C^* \right\}.$$

(Hint: If $(\exp tA)' J(\exp tA) = J$ for all t, differentiate and evaluate at $t = 0$ to obtain $A^* J + JA = 0$. Multiply this out to get the preceding

conditions. Conversely any block matrix satisfying the conditions can be written as

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & -A^* \end{bmatrix} + \begin{bmatrix} 0 & B \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ C & 0 \end{bmatrix}.$$

Show directly that each of the summands is in $\mathfrak{sp}(V)$ and use the fact that $\mathrm{sp}(V)$ is a subspace.)

We introduce another general technique, this time one that applies to semigroups and groups.

Proposition 23. *Let W be a closed convex cone in the vector space* End *V that is also closed under multiplication. Then $S := (I + W) \cap \mathrm{GL}(V)$ is a closed subsemigroup of $\mathrm{GL}(V)$ and $L(S) = W$.*

Proof. Let $X, Y \in W$. Then $(I + X)(I + Y) = I + X + Y + XY \in I + W$ since W is closed under multiplication and addition. Thus $I + W$ is a closed subsemigroup, and thus its intersection with $\mathrm{GL}(V)$ is a subsemigroup closed in $\mathrm{GL}(V)$.

Let $A \in W$. Then for $t \geq 0, \exp(tA) = I + \sum_{n=1}^{\infty} t^n A^n / n!$ has all finite partial sums in $I + W$ since W is closed under multiplication, addition, and scalar multiplication. Since the whole sum is the limit, it follows that $\exp(tA)$ is in the closed set $I + W$, and since the exponential image is invertible, it is in S. Thus $A \in L(S)$.

Conversely assume that $\exp(tA) \in S$ for all $t \geq 0$. Then

$$A = \frac{d}{dt}\big|_{t=0} \exp(tA) = \lim_{t \to 0^+} \frac{\exp(tA) - I}{t} \in W,$$

where the last assertion follows from the fact that $\exp(tA) \in I + W$ for $t > 0$, and hence $\exp(tA) - I$ and therefore $(1/t)(\exp(tA) - I)$ are in W. Since W is closed the limit is also in W.

Exercise 10.4. Use Proposition 23 to show the following in $\mathrm{GL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{C})$.

i. The group of unipotent (diagonal entries all 1) upper triangular matrices has Lie algebra the set of strictly upper triangular matrices.

ii. The group of invertible upper triangular matrices has Lie algebra the set of all upper triangular matrices.

iii. The group of stochastic matrices (invertible matrices with all row sums 1) has Lie algebra the set of matrices with all row sums 0.

iv. The semigroup S of all invertible matrices with all entries non-negative has as its Lie wedge the set of matrices whose nondiagonal entries are nonnegative.

Suggestions for Further Reading

The author is indebted to two sources in particular that have greatly influenced these notes [1-6]. The first is Roger Howe's "Very Basic Lie Theory" [6]. Comparison with that source will show our indebtedness to that source, particularly in treatment of matters connected with the Baker-Campbell-Hausdorff formula. A second important source is what arose as the class notes on introductory Lie theory by Karl Hofmann, much of which was has been incorporated in Chapter 5 of the book "The Structure of Compact Groups" [5]. The idea of introducing Lie theory via matrix Lie theory also has been worked out by Morton Curtis [1]. A novel feature of these notes is the inclusion toward the end of the fairly recent notion of a Lie semigroup. More on this topic

can be found in the literature [2,3]. For a comprehensive treatment of Lie theory, one somewhat in the spirit of these notes, especially in its earlier parts, the author recommends the recent volume of Hilgert and Neeb [4].

References

1. Curtis M (1984) Matrix Groups, Second edition, Universitext, Springer-Verlag, New York.

2. Hilgert J, Hofmann KH, Lawson JD (1989) Lie Groups, Convex Cones, and Semigroups, Oxford Mathematical Monographs, Oxford Science Publications, New York.

3. Hilgert J, Neeb KH (1993) Lie Semigroups and Their Applications. Lecture Notes in Mathematics, Springer-Verlag, Berlin.

4. Hilgert J, Neeb KH (2012) Structure and Geometry of Lie Groups. Springer Monographs in Mathematics, Springer, New York.

5. Hofmann KH, Morris SA (2013) The Structure of Compact Groups: A primer for the student - A handbook for the expert. In: De Gruyter Berlin, Revised and Augmented (3rdedn), De Gruyter Studies in Mathematics.

6. Howe R (1983) Very basic Lie theory. Amer. Math. Monthly 90: 600-623.