Lie Algebras
and their corresponding linear groups

Betsy Kersey
MATH 434
University of Puget Sound

Spring 2010

1 Some Linear Algebra Review

In the study of Lie algebras and linear groups, there are several topics from linear algebra that play an important role in the definitions and theorems. They are presented here as a convenience, so that you will not have to dig out your old linear algebra book and look them up yourself.

Definition 1.1 *The* trace *of a matrix A, $tr(A)$, is the sum of the entries along the diagonal of the matrix A.*

Definition 1.2 *The* norm *of a matrix A is the number identified by $\|A\| = max\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$, where $\mathbf{x}$ is any vector other than the zero vector. The norm is also known as the "amplifying power" of the matrix.*

The following theorems and definitions will give us a better idea as to how one may compute the norm of a matrix.

Definition 1.3 *The conjugate transpose, or* adjoint*, of an m x n matrix A is the n x m A^* such that $(A^*)_{ij} = \overline{A_{ji}}$. A matrix is* self-adjoint *if $A^* = A$.*

Definition 1.4 *Let B be an n x n self-adjoint matrix. The* Rayleigh quotient *for $\mathbf{x} \neq \mathbf{0}$ is defined as the scalar $R_B(\mathbf{x}) = \frac{(B\mathbf{x},\mathbf{x})}{\|\mathbf{x}\|^2}$, where $(B\mathbf{x},\mathbf{x})$ is the inner product of $B\mathbf{x}$ and $\mathbf{x}$.*

Theorem 1.5 *For a self-adjoint matrix B, the max of the Rayleigh quotient when $\mathbf{x} \neq \mathbf{0}$, $max_{\mathbf{x}\neq\mathbf{0}}R_B(\mathbf{x})$, is equal to the largest eigenvalue of B, and the min of the Rayleigh quotient when $\mathbf{x} \neq \mathbf{0}$, $min_{\mathbf{x}\neq\mathbf{0}}R_B(\mathbf{x})$, is equal to the smallest eigenvalue of B.*

Proof There is an earlier theorem stating that if V is a finite-dimensional vector space, and T is a linear transformation in V, then T is self-adjoint if and only if V has an orthonormal basis consisting of eigenvectors of V. By this theorem, since we know B is self-adjoint, we can choose an orthonormal basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ of eigenvectors of B such that $B\mathbf{x}_i = \lambda_i\mathbf{x}_i$, $1 \leq i \leq n$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Since the eigenvalues of a self-adjoint matrix are always real, these inequalities are meaningful. Now for $\mathbf{x} \in \mathbf{C}^n$, there exist scalars $a_1, \ldots, a_n$ such that

$$\mathbf{x} = \sum_{i=1}^{n} a_i\mathbf{x}_i$$

and thus

$$R_B(\mathbf{x}) = \frac{(B\mathbf{x},\mathbf{x})}{\|\mathbf{x}\|^2} = \frac{(\sum_{i=1}^{n}\lambda_i a_i\mathbf{x}_i, \sum_{j=i}^{n} a_j\mathbf{x}_j)}{\|\mathbf{x}\|^2} = \frac{\sum_{i=1}^{n}\lambda_i|a_i|^2}{\|\mathbf{x}\|^2} \leq \frac{\lambda_1\sum_{i=1}^{n}|a_i|^2}{\|\mathbf{x}\|^2} = \lambda_1.$$

It should be obvious that $R_B(\mathbf{x}_1) = \lambda_1$, and so we have shown that $max_{\mathbf{x}\neq\mathbf{0}}R_B(\mathbf{x})$ is equal to the largest eigenvalue of B. The second part of the theorem, that $min_{\mathbf{x}\neq\mathbf{0}}R_B(\mathbf{x})$ is equal to the smallest eigenvalue of B, is proved in a similar fashion.[2] ∎

Corollary 1.6 *For any square matrix A, $\|A\|$ is finite and is equal to $\sqrt{\lambda}$, where λ is the largest eigenvalue of A^*A.*

Proof Let B be the self-adjoint matrix A^*A (since $(A^*A)^* = A^*(A^*)^* = A^*A$), and let λ be the largest eigenvalue of B. Since when $x \neq 0$, we have

$$0 \leq \frac{\|Ax\|^2}{\|x\|^2} = \frac{(Ax, Ax)}{\|x\|^2} = \frac{(A^*Ax, x)}{\|x\|^2} = \frac{(Bx, x)}{\|x\|^2} = R_B(x),$$

the preceeding theorem tells us that $\|A\|^2 = \lambda$. ∎

2 The Exponential of a Matrix

Unless otherwise specified, all matrices are taken to be square.

Definition 2.1 *If X is a matrix, then we define* $\exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} X^k$.

This should look familiar, as this is also the Taylor series for e^x when x is a real or complex variable. This definition is given further justification in the following proposition.

Proposition 2.2 *a) For any matrix X, $\frac{d}{d\tau} \exp(\tau X) = X \exp(\tau X) = \exp(\tau X) X$, and $a(\tau) = \exp(\tau X)$ is the unique differentiable solution to $a'(\tau) = Xa(\tau)$, $a(0) = 1$.*

b) For any two commuting matrices X and Y, $\exp(X)\exp(Y) = \exp(X + Y)$.

c) For any matrix X, $\exp(X)$ is invertible, and $(\exp(X))^{-1} = \exp(-X)$.

d) For any matrix X and scalars τ, σ, $\exp((\tau + \sigma)X) = \exp(\tau X)\exp(\sigma X)$, and $a(\tau) = \exp(\tau X)$ is the unique differentiable solution of $a(\tau+\sigma) = a(\tau)a(\sigma)$, $a(0) = 1$, and $a'(0) = X$.

Proof a) Let us take the definition of $\exp(\tau X)$ and differentiate it term-by-term:

$$\frac{d}{d\tau}(\exp(\tau X) = \frac{d}{d\tau} \sum_{j=0}^{\infty} \frac{\tau^j}{j!} X^j = \sum_{j=1}^{\infty} \frac{\tau^{j-1}}{(j-1)!} X^j.$$

Factoring out an X from this series, either to the left or to the right, and then reindexing shows that this is equal to $X \exp(\tau X) = \exp(\tau X) X$. For the second claim of a), assume that $a(\tau)$ satisfies the equations $a'(\tau) = Xa(\tau)$ and $a(0) = 1$. Differentiating according to the product rule, we get

$$\begin{aligned} \frac{d}{d\tau}(\exp(-\tau X)(a(\tau)) &= (\frac{d}{d\tau}\exp(-\tau X))a(\tau) + \exp(-\tau X)(\frac{d}{d\tau}a(\tau)) \\ &= \exp(-\tau X)(-X)a(\tau) + \exp(-\tau X)Xa(\tau) \\ &= 0. \end{aligned}$$

Thus, $\exp(-\tau X)a(\tau)$ is independent of τ, and since this equation is 1 at $\tau = 0$, it equals one identically. Now part c) tells us that this means $\exp(\tau X)^{-1}a(\tau) = 1$, and thus $a(\tau) = \exp(\tau X)$.

b) The product of two norm-convergent matrix series can be calculated by forming all possible products of terms of the first series with terms of the second series and then summing in any order. When we apply this principle to $\exp X \exp Y$, we get

$$\exp X \exp Y = (\sum_{j=0}^{\infty} \frac{X^j}{j!})(\sum_{k=0}^{\infty} \frac{Y^k}{k!}) = \sum_{j,k=0}^{\infty} \frac{X^j Y^k}{j!k!}.$$

Considering the other side of the equation, and assuming that X and Y commute, we can rearrange and collect terms while expanding $(1/m!)(X+Y)^m$ to get

$$\exp(X+Y)=\sum_{m=0}^{\infty}\frac{1}{m!}(X+Y)^m=\sum_{m=0}^{\infty}\frac{1}{m!}(\sum_{j+k=m}\frac{m!}{j!k!}X^jY^k)=\sum_{j,k=0}^{\infty}\frac{X^jY^k}{j!k!}.$$

Since this is equal to what we obtained from the left side of the equation, we conclude that $\exp(X)\exp(Y)=\exp(X+Y)$.

c) We must show that $\exp(X)\exp(-X)=1$. This is done if we take b) and put in $Y=-X$. Then we have $\exp(X)\exp(-X)=\exp(X-X)=\exp(0)=1$. This assumes that X and $-X$ commute, which is easily shown to be true.

d) If we distribute the scalars, then we have $\exp((\tau+\sigma)X)=\exp(\tau X+\sigma X)$, and with b), this is easily shown to be equal to $\exp(\tau X)(\exp(\sigma X)$. For the second claim of d), assume that $a(\tau)$ has the indicated property. Then

$$a'(\tau)=\frac{d}{d\sigma}a(\tau+\sigma)|_{\sigma=0}=a(\tau)\frac{d}{d\sigma}|_{\sigma=0}=a(\tau)X,$$

and the claim follows from the second part of a). [3] ∎

Proposition 2.3 *For any matrix X and any invertible matrix A, $A\exp(X)A^{-1}=\exp(AXA^{-1})$.*

Proof

$$A\exp(X)A^{-1}=A(\sum_{k=0}^{\infty}\frac{X^k}{k!})A^{-1}=\sum_{k=0}^{\infty}\frac{(AXA^{-1})^k}{k!}=\exp(AXA^{-1}).\quad\blacksquare$$

Definition 2.4 *For any invertible matrix $A\in M$, where M is a matrix space, we denote by Ad_A the conjugation operation by A as a linear transformation of M: $Ad_A(Y)=AYA^{-1}$. Conjugation is here used in the sense we meant it in group theory last semester, not as it would mean in regard to complex numbers. Note that $Ad_A(XY)=Ad_A(X)Ad_A(Y)$, and $Ad_A(X^{-1})=Ad_A(X)^{-1}$, so that Ad_A is a homomorphism from $M\to GL(M)$.*

Definition 2.5 *For a not necessarily invertible matrix $X\in M$, we define the operation ad_X as $ad_X(Y)=XY-YX$. ad_X is a linear transformation of M. We generally write this as a* bracket*: $[X,Y]=XY-YX$. This is also known as the* commutator*.*

The following proposition shows the connection between Ad and ad.

Proposition 2.6 *For any $X\in M$, $Ad_{\exp X}(Y)=\exp(ad_X(Y))$.*

Proof Fix $X\in M$ and let $A(\tau)=Ad_{\exp(\tau X)}Y$ for $\tau\in\mathbf{R}$. Calculate the derivative:

$$\begin{aligned}A'(\tau)&=\frac{d}{d\tau}(\exp(\tau X)Y\exp(-\tau X))\\&=X\exp(\tau X)Y\exp(-\tau X)+\exp(\tau X)Y\exp(-\tau X)(-X)\\&=ad_X(Ad_{\exp(\tau X)}(Y)).\end{aligned}$$

Thus, $A'(\tau)=ad_X(A(\tau))$, and $A(0)=1$. From Proposition 2.1, we know that the only solution of these equations is $A(\tau)=\exp(\tau ad_X(Y))$, which gives the desired result. [3] ∎

Theorem 2.7 *Let $X = X(\tau)$ be a matrix-valued differentiable function of a scalar variable τ. Then*

$$\frac{d}{d\tau}\exp(X) = \exp(X)\frac{1-\exp(-ad_X)}{ad_X}\frac{dX}{d\tau}.$$

Proof Set $Y(\sigma,\tau) = \exp(-\sigma X(\tau))\frac{\partial}{\partial\tau}\exp(\sigma X(\tau))$ for $\sigma,\tau \in \mathbf{R}$. Differentiate with respect to σ:

$$\begin{aligned}
\frac{\partial Y}{\partial\sigma} &= (\exp(-\sigma X))(-X)\frac{\partial}{\partial\tau}\exp(\sigma X) + \exp(-\sigma X)\frac{\partial}{\partial\tau}(X\exp(\sigma X))\\
&= (\exp(-\sigma X))(-X)\frac{\partial}{\partial\tau}\exp(\sigma X) + \exp(\sigma X)\frac{dX}{x\tau}\exp(\sigma X) + (\exp(-\sigma X))X\frac{\partial}{\partial\tau}\exp(\sigma X)\\
&= (\exp(-\sigma X))\frac{dX}{d\tau}\exp(\sigma X)\\
&= Ad_{\exp(-\sigma X)}\frac{dX}{d\tau}\\
&= \exp(ad_{-\sigma X}(\frac{dX}{d\tau})).
\end{aligned}$$

Now

$$\exp(-X)\frac{d}{d\tau}\exp X = Y(1,\tau) = \int_0^1 \frac{\partial}{\partial\sigma}Y(\sigma,\tau)\,d\sigma[since Y(0,\tau)=0]$$

and

$$\frac{\partial Y}{\partial\sigma} = \exp(-\sigma ad_X(\frac{dX}{d\tau})) = \sum_{k=0}^{\infty}\frac{(-1)^k\sigma^k}{k!}(ad_X)^k\frac{dX}{d\tau}.$$

When we integrate this series term by term from $\sigma = 0$ to $\sigma = 1$, we get

$$\exp(-X)\frac{d}{d\tau}\exp X = \sum_{k=0}^{\infty}\frac{(-1)^k}{(k+1)!}(ad_X)^k\frac{dX}{d\tau},$$

which, with a little manipulation, is the formula we were after. [3] ∎

Lemma 2.8 *Let $f(z) = \sum_{k=0}^{\infty}\alpha_k z^k$ be a power series with real or complex coefficients. Suppose U is a linear transformation so that the series $f(U) = \sum_{k=0}^{\infty}\alpha^k U^k$ converges. If $\lambda_1,\ldots,\lambda_n$ are the eigenvalues of U, listed with multiplicities, then $f(\lambda_1),\ldots,f(\lambda_n)$ are the eigenvalues of $f(U)$, listed with multiplicities.*

Proof Select a basis so that U is triangular with diagonal entries $\lambda_1,\ldots,\lambda_n$. (This may require going to complex scalars, even if the original matrix has only real coefficients.) For each $k = 0,1,\ldots,U^k$ is then also triangular with diagonal entries $\lambda_1^k,\lambda_2^k,\ldots,\lambda_n^k$. Thus, $f(U)$ itself is a triangular matrix with diagonal entries given by the power series $f(\lambda_1), f(\lambda_2),\ldots,f(\lambda_n)$; these power series $f(\lambda_j)$ converge, because $f(U)$ is assumed to converge. ∎

Proposition 2.9 *If none of the eigenvalues of ad_X are of the form $\lambda = 2\pi ik, k = \pm1,\pm2,\ldots,$ then $\exp : M \to M$ has an inverse defined on a neighborhood of X.*

Proof By the above lemma, the eigenvalues of $(1-\exp(-U))/U$ are of the form $(1-e^{-\lambda})/\lambda$, λ an eigenvalue of U. The given values of λ are precisely the solutions of the equation $(1-e^{-\lambda})/\lambda = 0$; this gives the conclusion when one takes $U = ad_X$. ∎

3 Lie Algebras and Linear Groups as Examples

Through this section, and indeed throughout the entire paper, "Lie" should be pronounced "lee," not "lye." We begin with the definition of a Lie algebra.

Definition 3.1 *A* Lie algebra g *is a real (or complex) vector space* g *over a field* F *together with an operator* $[,] : g \times g \to g$ *called the* bracket*, such that for* $x, y, z \in g$*, the following properties hold:*

1) $[x, y] \in g$ *(closure)*

2) $[ax + by, z] = a[x, z] + b[y, z]$ *and* $[z, ax + by] = a[z, x] + b[z, y]$ *for all scalars* $a, b \in F$ *(bilinearity)*

3) $[x, y] = -[y, x]$ *(anticommutativity)*

4) $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0.$ *(Jacobi identity)*

Example Any vector space V, if $[x, y] = 0$ for all $x, y \in V$, is a Lie algebra. Such an algebra is called *abelian*. This is consistent with our former definition of abelian as commutative, since by property (2) of the definition, the only way we can have $[x, y] = [y, x]$ is if $[y, x] = -[y, x]$, so that $[y, x] = 0$ for all x, y.

Example The set of vectors $\mathbf{R}^3$ is a Lie algebra under the operation $[X, Y] = X \text{ x } Y$, where $X \text{ x } Y$ is the cross-product.

Example Any vector space with dimension 2 and basis x, y is a Lie algebra if we set $[x, x] = [y, y] = 0$ and $[x, y] = y$, and then extend these properties so that they satisfy the bilinearity property.

Example The vector space of all real n x n matrices is a Lie algebra if $[A, B] = AB - BA$. Since this is one of the most important examples, we will prove that this satisifies all four properties of the definition.

Closure: Since the vector space of all real n x n matrices is closed under matrix multiplication and matrix addition (and therefore subtraction), it is also closed under the bracket operation.

Bilinearity:

$$\begin{aligned}
[\alpha A + \beta B, C] &= (\alpha A + \beta B)C - C(\alpha A + \beta B) \\
&= \alpha AC + \beta BC - C\alpha A - C\beta B \\
&= \alpha AC - \alpha CA + \beta BC - \beta CB \\
&= \alpha(AC - CA) + \beta(BC - CB) \\
&= \alpha[A, C] + \beta[B, C]
\end{aligned}$$

$$\begin{aligned}
[C, \alpha A + \beta B] &= C(\alpha A + \beta B) - (\alpha A + \beta B)C \\
&= -((\alpha A + \beta B)C - C(\alpha A - \beta B)) \\
&= -[\alpha A + \beta B, C] \\
&= -(\alpha[A, C] + \beta[B, C]) \\
&= -\alpha(AC - CA) - \beta(BC - CB) \\
&= \alpha(CA - AC) + \beta(CB - BC) \\
&= \alpha[C, A] + \beta[C, B]
\end{aligned}$$

Anticommutativity: $[A,B] = AB - BA = -(BA - AB) = -[B,A]$.
Jacobi identity:

$$\begin{aligned}
[[A,B],C]+[[B,C],A]+[[C,A],B] &= [AB-BA,C]+[BC-CB,A]+[CA-AC,B] \\
&= (AB-BA)C - C(AB-BA) + (BC-CB)A \\
&\quad -A(BC-CB) + (CA-AC)B - B(CA-AC) \\
&= ABC - BAC - CAB + CBA + BCA - CBA \\
&\quad -ABC + ACB + CAB - ACB - BCA + BAC \\
&= 0.
\end{aligned}$$

Therefore, this truly is a Lie algebra, since it satisfies all of the requisite four properties. You will notice that this is the same bracket operation we defined earlier, which is now justified. Note that it was also referred to as $ad_X(Y)$, in addition to as $[X,Y]$. ∎

Definition 3.2 *A* linear group *G is any family of invertible matrices with the property that $A \in G$ implies that $A^{-1} \in G$, and $A, B \in G$ implies that $AB \in G$.*

Definition 3.3 *A* linear Lie algebra *g is a Lie algebra over a linear group G with the bilinear operation defined by $[A,B] = AB - BA$.*

Definition 3.4 *Let G be a linear group. The* tangent space *g to G at 1 consists of all matrices X for which one can find a differentiable curve $a(\tau)$ that lies in G and satisfies $a(0) = 1$ and $a'(0) = X$.*

Proposition 3.5 *1) g, the tangent space to a linear group, is a real vector space:*

$$X, Y \in g \text{ implies } \alpha X + \beta Y \in g \text{ for all } \alpha, \beta \in \mathbf{R}.$$

2) g is closed under the bracket operation:

$$X, Y \in g \text{ implies } [X,Y] \in g.$$

Proof 1) Take $X, Y \in g$ and $\alpha, \beta \in \mathbf{R}$. There are differentiable curves $a(\tau)$ and $b(\tau)$ that lie in G so that $a(0) = b(0) = 1$, $a'(0) = X$, and $b'(0) = Y$. Then the differentiable curve $c(\tau) = a(\alpha\tau)b(\beta\tau)$ also lies in G. It satisfies $c(0) = 1$ and

$$c'(0) = (a'(\alpha\tau)\alpha b(\tau) + a(\alpha\tau)b'(\beta\tau)\beta)|_{\tau=0} = \alpha X + \beta Y.$$

This shows that $\alpha X + \beta Y \in g$ as well. 2) Again take $X, Y \in g$ and choose $a(\tau)$ and $b(\tau)$ as above. For fixed $\sigma \in \mathbf{R}$ (sufficiently close to 0), $c_\sigma(\tau) = a(\sigma)b(\tau)a(\sigma)^{-1}$ is a differentiable curve in G. It satisfies $c_\sigma(0) = 1$ and $c'_\sigma(0) = a(\sigma)Ya(\sigma)^{-1}$. This shows that $a(\sigma)Ya(\sigma)^{-1} \in g$ for σ sufficiently close to 0. As a curve in the vector space g, $\sigma \to a(\sigma)Ya(\sigma)^{-1}$ has its tangent vector in g. We compute the following:

$$\frac{d}{d\sigma}(a(\sigma)Ya(\sigma)^{-1})|_{\sigma=0} = XY - YX = [X,Y].$$

This shows that $[X,Y] \in g$, as required.[3] ∎

From the above, it should be evident that the tangent space of G and the Lie algebra of G are really the same thing.

Theorem 3.6 *If $f : \mathbf{R}^n \to \mathbf{R}^n$ is a k times differentiable map, $1 \leq k \leq \infty$, with an invertible derivative at a point $c \in \mathbf{R}^n$, then f is itself k times differentiable and invertible in a neighborhood of c.*

This is known as the Inverse Function Theorem. Since it belongs more in the reign of calculus, we shall not prove it here, but go on to use it in the following theorem.

Theorem 3.7 *Let G be a linear group, and g its Lie algebra. Then* exp *maps g into G.*

Proof Choose a basis $X_1, X_2, \ldots, X_m$ for g and set

$$f(\tau_1 X_1 + \tau_2 X_2 + \ldots + \tau_m X_m) = a_1(\tau_1)a_2(\tau_2)\cdots a_m(\tau_m),$$

where $a_k(\tau)$ is a differentiable curve with $a_k(0) = 1$ and $a_k'(0) = X_k$. Then $f : g \to G \subset M$ and $df_0 X = X$ for all $X \in g$. Choose a subspace s of M complementary to g, so that $M = g \oplus s$ (the direct sum of g and s), and a map $h : s \to M$ partially defined in a neighborhood of 0 in s and satisfying $h(0) = 1$ and $dh_0 Y = Y$ for all $Y \in s$; for example, $h(Y) = 1 + Y$. Define $\phi : g \times s \to M$ by $\phi(X, Y) = f(X)h(Y)$. Then ϕ is defined and differentiable in a neighborhood of (0,0) in $g \times s$ and satisfies $df_{(0,0)}(X, Y) = X + Y$. In particular, this differential is invertible.

According to the Inverse Function Theorem, ϕ itself has a local inverse $W : M \to g \times s, a \to (U(\alpha), V(\alpha))$, defined for a in a neighborhood of 1 in M is of the form $a = f(X)h(Y)$ for unique (X, Y) in a neighborhood of (0,0) in $g \times s$, namely $X = U(a), Y = V(a)$. If $V(a) = 0$, then $a = f(X)h(0) = f(X)$ is in G. It is also the case that $V(a) = 0$ implies that $a \in G$.

Fix $(X, Y) \in g \times s$, sufficiently close to (0,0) so that W is defined near $a = \phi(X, Y)$ in M. Given $Z \in g$, we have

$$V(f(X + \tau Z)h(Y)) = Y$$

for $\tau \in \mathbf{R}$ close to 0. Differentiation with respect to τ at $\tau = 0$ gives

$$(dV)_a\{(df_X Z)(h(Y)\} = 0.$$

Substituting $h(Y) = f(X)^{-1}a$, we obtain

$$dV_a\{(df_X Z)f(X)^{-1}a\} = 0.$$

The matrix $(df_X Z)f(X)^{-1}$ is again in g, because it is the tangent vector at 1 of a differentiable curve in G:

$$((df)_X Z)g(X)^{-1} = \frac{d}{d\tau} f(X + \tau Z)f(X)^{-1}|_{\tau=0}.$$

Thus, we can define a map $A : g \to g$ by $AZ = (df_X Z)f(X)^{-1}$. $A = A(X)$ is a linear transformation of g depending continuously on X. $dV_a\{(df_X Z)f(X)^{-1}a\} = 0$ now becomes

$$dV_a\{(AZ)a\} = 0,$$

where it is understood that

$$a = f(X)h(X), A = A(X).$$

For $X = 0, A$ reduces to the identity transformation of g, because $f(0) = 1$ and $(df)_0 Z \equiv Z$. In particular, $\det A \neq 0$ for $X = 0$. By continuity, $\det A \neq 0$ for X in a neighborhood of 0 in g. So if X is near 0 in g, then A is invertible, and every element of g is of the form AZ for some $Z \in g$. We can therefore replace AZ by Z in $dV_a\{(AZ)a\} = 0$ and see that

$$dV_a(Za) = 0$$

for all a in a neighborhood of 1 in M and all Z in g.

We now return to the assertion that exp maps g into G. Take $X \in g$ and set $a(\tau) = \exp(\tau X)$. According to $dV_a(Za) = 0$,

$$dV_{a(\tau)}(Xa(\tau)) = 0,$$

which says that

$$\frac{d}{d\tau} V(a(\tau)) = 0;$$

and this holds for all τ in an interval about 0 in $\mathbf{R}$. Consequently, $V(a(\tau))$ is constant there. For $\tau = 0$ we have $V(a(0)) = 0$, and therefore $V(a(\tau)) \equiv 0$ for τ in an interval about 0 in $\mathbf{R}$. When we consider that $V(a) = 0$ implies that $a \in G$, this implies that $a(\tau) = exp(\tau X)$ lies in G for τ in that interval. Contemplating the identity $\exp X = (\exp X/k)^k$, one sees that $\exp(X)$ will be in G for *all* $X \in g$. [3] ∎

Corollary 3.8 *The Lie algebra g of a linear group G consists of all matrices X for which $\exp(\tau X)$ lies in G for all $\tau \in \mathbf{R}$.*

Proof Let us use the notation $S = \{X | \exp(\tau X) \subset G \forall \tau \in \mathbf{R}\}$. The fact that $g \subset S$ follows from the above theorem. The fact that $S \subset g$ follows when one observes that $a(\tau) = \exp(\tau X)$ provides a path for X, as required by the definition of g as a tangent space.

4 Properties of Lie Algebras

There are many structures of algebras we have seen earlier in our course that have similar parallels in Lie algebras. We list several of them here, although there is not really space to do them justice.

Definition 4.1 *A* homomorphism of algebras *$\phi : g \to h$ of Lie algebras is a linear map that respects brackets: $\phi([X, Y]) = [\phi(X), \phi(Y)]$. An* isomorphism *is an invertible homomorphism. An isomorphism $\alpha : g \to g$ of a Lie algebra g with itself is called an* automorphism *of g. The collection of all automorphisms of a Lie algebra g is a subgroup of $GL(g)$ (the group of linear transformations of g) and therefore, a linear group. It is denoted $Aut(g)$.*

Definition 4.2 *When g is the Lie algebra of a linear group G, then every $A \in G$ defines an automorphism $Ad_A(X) = AXA^{-1}$. The map $G \to Aut(g)$, $A \to Ad_A$, is a group homomorphism, and is called the* adjoint representation*.*

Definition 4.3 *A* Lie subalgebra *of g is a subspace closed under the bracket operation. A Lie subalgebra a is an* ideal *if $[A, B] \in a$ whenever $A \in a$ and $B \in g$. If g has no ideals other than $\{0\}$ and g, then g is called a* simple Lie algebra.

Definition 4.4 *The* center *of g is the largest ideal c such that $[g, c] = 0$; i.e., that $[X, Y] = 0$ for all $X \in g$, $Y \in c$. It is unique.*

Definition 4.5 *Any linear transformation δ of g with the property $\delta([Y, Z]) = [\delta(Y), Z] + [Y, \delta(Z)]$ is called a* derivation *of g. The collection of all derivations of a Lie algebra g, denoted by $Der(g)$, is a Lie subalgebra of $gl(g)$, and therefore, a linear Lie algebra.*

Proposition 4.6 *$Der(g)$ is the Lie algebra of the linear group $Aut(g)$.*

Proof Using Corollary 3.4, we have to show that a linear transformation δ of g is a derivation, that is, it satisfies

$$\delta([X, Y]) = [\delta(X), Y] + [X, \delta(Y)]$$

if and only if $exp(\tau\delta)$ is an automorphism for all $\tau \in \mathbf{R}$, i.e. it satisfies

$$exp(\tau\delta)[X, Y] = [exp(\tau\delta)X, exp(\tau\delta)Y].$$

If the latter equation holds for all τ, then then the former equation follows by differentiation at $\tau = 0$, using the product rule for bilinear mappings.

Conversely, suppose the former equation holds. To see that the latter equation holds as well, write it in the form

$$exp(-\tau\delta)[exp(\tau\delta)X, exp(\tau\delta)Y] = [X, Y].$$

The derivative of the left side of this equation with respect to τ is

$$exp(-\tau\delta)(-\delta[exp(\tau\delta)X, exp(\tau\delta Y] + [\delta exp(\tau\delta)X, exp(\tau\delta)Y] + [exp(\tau\delta)X, \delta exp(\tau\delta)Y]),$$

which is identically 0, in view of our first equation. Thus the left side that we just differentiated is independent of τ, and therefore equal to the right side, as one sees by setting $\tau = 0$ [3] ∎

5 Resources

References

[1] Belinfante, Johan G.F. and Bernard Kolman. *A Survey of Lie Groups and Lie Algebras.* Philadelphia: Society for Industrial and Applied Mathematics, 1972.

[2] Friedberg, Stephen H., Arnold J. Insel, and Lawrence E. Spence. *Linear Algebra.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

[3] Rossman, Wulf. *Lie Groups: An Introduction Through Linear Groups.* Oxford: Oxford University Press, 2002.

[4] Sattinger, D.H. and O.L. Weaver. *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*. New York: Springer-Verlag, 1986.

[5] Strang, Gilbert. *Linear Algebra and its Applications.* New York: Academic Press, 1976.

[6] Varadarajan, V.S. *Lie Groups, Lie Algebras, and Their Representations*. New York: Springer-Verlag, 1974.

[7] Warner, Frank W. *Foundations of Differentiable Manifolds and Lie Groups.* Glenview, IL: Scott, Foresman and Company, 1971.