An Overview of Lie Groups and Algebras

Larry Susanka

November 28, 2023

An Overview of Lie Groups and Algebras

Table of Contents

Introduction

Groups of Square Matrices

Algebras of Square Matrices

The Exponential and Logarithm

Lie Algebra ⇔ Lie Group

Introduction

Here we take a quick look at Lie groups and algebras. These were invented around and after 1873 by Norwegian mathematician Sophus Lie, giving specific content and methods to the Erlangen program of Felix Klein, announced the previous year.

Wilhelm Killing made major progress in a series of papers dating from 1888. Many of the properties of these groups and algebras were explored in the next few decades.

A small zoo of ten or twenty of these has been collected and curated by particle physicists and relativists, introduced to these folks by Hermann Weyl in 1927. Understanding how these behave is not optional if you wish to read (and understand) almost any modern work on theoretical physics.

Schrödinger himself was reported to have lamented this “gruppenpest,” a term invented by disgruntled physicists of the day who found early treatments to be opaque. But lately physicists have just accepted, without a lot of fuss, that these must be included in their language.

This talk will provide the most basic facts about matrix Lie groups and algebras, to give a sense of the issues in the area. Anyone who wants to explore further has numerous options.

You can’t go wrong with John Stillwell’s excellent and short *Naive Lie Theory*, which includes extensive discussion of the matrix groups most often encountered as well as historical overview of this important subject and a wonderful variety of results proved by elementary means.

Stillwell includes a short purely algebraic proof of the **Baker-Campbell-Hausdorff formula**. This proof was discovered by M. Eichler in 1968 and published as *A New Proof of the Baker-Campbell-Hausdorff Formula*.

Defining the Lie bracket (also known as the commutator or Lie product) of two square matrices X and Y by $[X, Y] = XY - YX$ the **BCH formula** states that if X and Y are sufficiently small square matrices then a specific series of the form

$$X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] + \frac{1}{12}[Y,[X,Y]] + \text{ higher order commutators of } X \text{ and } Y$$

converges to a matrix Z and $e^X e^Y = e^Z$.

At a slightly more advanced (but still very accessible) level I recommend Brian C. Hall's *Lie Groups, Lie Algebras, and Representations*.

To understand why the BCH formula is important we need to talk a bit about matrix groups and algebras of square matrices and the relationship between these two.

In this talk ALL groups will be groups of square matrices with matrix multiplication and ALL algebras (even if they have complex entries) will be **real** vector spaces of matrices.

These sets will be subsets of $\mathbb{M}_{n\times n}(\mathbb{R})$ or $\mathbb{M}_{n\times n}(\mathbb{C})$, the square matrices with either real or complex entries.

But I get ahead of myself here: we need to define what we are talking about first.

Groups

A matrix group G is, first of all, a set of $n \times n$ matrices. These matrices might be real, or they could have complex entries. This collection of invertible matrices must contain the multiplicative identity matrix I_n and whenever $M, N \in G$ we also have $M^{-1} \in G$ and $MN \in G$.

That is, the collection is closed under matrix multiplication and the taking of inverses.

A matrix Lie group has one additional property:

If M_i is a sequence of members of G and M_i converges[1] to A then either $A \in G$ or A is not invertible.

[1] $M_i \to A$ means all n^2 entries converge to the corresponding entry of A.

The smallest example of a Lie group is the 1×1 matrix group consisting of $\{1\}$. Not very interesting, but how about $\{-1, 1\}$? This is *slightly* more interesting.

$\mathbb{R}_{\emptyset}$ and $\mathbb{R}_+$ consisting of the nonzero real numbers and the positive real numbers are two more.

You may recall the complex numbers as the set of symbols "of the form $a + bi$" where a and b are real and $i^2 = -1$. A complex number is "identified" with a point (a, b) in the plane, a 2-dimensional real vector space.

Any such point has a "polar form" $\sqrt{a^2+b^2}(\cos(\theta), \sin(\theta))$ where θ is the angle of the point measured counterclockwise from the positive x axis and complex multiplication with $(c, d) = \sqrt{c^2+d^2}(\cos(\mu), \sin(\mu))$ is calculated as

$$\sqrt{a^2+b^2}\sqrt{c^2+d^2}(\cos(\theta+\mu), \sin(\theta+\mu)).$$

The complex numbers can also be construed as a vector space of matrices of the form

$$a\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + b\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}a & -b\\ b & a\end{pmatrix}.$$

Matrix multiplication matches the multiplication pattern for complex numbers.

The determinant is the squared norm of the complex number, and the transpose is the complex conjugate.

$\mathbb{C}_\emptyset$ and $\mathbb{S}^1$ consisting of the nonzero complex numbers and the unit circle of complex numbers, the complex numbers of length 1, are two more Lie groups.

Define matrices

$$\boldsymbol{I_2} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \quad \boldsymbol{R} = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} \quad \boldsymbol{S} = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \quad \boldsymbol{T} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}.$$

and use these as 2×2 block matrices to form

$$w\begin{pmatrix}\boldsymbol{I_2} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I_2}\end{pmatrix} + x\begin{pmatrix}\boldsymbol{R} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{R}\end{pmatrix} + y\begin{pmatrix}\boldsymbol{0} & -\boldsymbol{S}\\ \boldsymbol{S} & \boldsymbol{0}\end{pmatrix} + z\begin{pmatrix}\boldsymbol{0} & -\boldsymbol{T}\\ \boldsymbol{T} & \boldsymbol{0}\end{pmatrix}$$

$$= w\boldsymbol{I_4} + x\boldsymbol{F} + y\boldsymbol{G} + z\boldsymbol{H} = \begin{pmatrix}w & -x & -y & -z\\ x & w & -z & y\\ y & z & w & -x\\ z & -y & x & w\end{pmatrix}.$$

A quaternion $w + x\,\vec{\boldsymbol{i}} + y\,\vec{\boldsymbol{j}} + z\,\vec{\boldsymbol{k}}$ may be identified with this matrix, and Hamilton product corresponds to matrix multiplication on this 4-dimensional real vector space.

$$w + x\,\vec{\boldsymbol{i}} + y\,\vec{\boldsymbol{j}} + z\,\vec{\boldsymbol{k}} \Longleftrightarrow \begin{pmatrix}w & -x & -y & -z\\ x & w & -z & y\\ y & z & w & -x\\ z & -y & x & w\end{pmatrix}$$

The magnitude of this quaternion is the 4th root of the determinant, and the quaternion conjugate is the transpose of this matrix.

$\mathbb{H}_\emptyset$ and $\mathbb{S}^3$, consisting of the nonzero quaternions and the unit sphere of unit quaternions, are two more Lie groups.

The **continuous Heisenberg group**, $H_3(\mathbb{R})$, is the set of 3×3 matrices of the form $\begin{pmatrix}1 & a & b\\ 0 & 1 & c\\ 0 & 0 & 1\end{pmatrix}$ with three real parameters.

$$\begin{pmatrix}1 & a & b\\ 0 & 1 & c\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\ t\\ 1\end{pmatrix} = \begin{pmatrix}x + at + b\\ t + c\\ 1\end{pmatrix}$$

so if $\begin{pmatrix}x\\ t\\ 1\end{pmatrix}$ is construed as representing a position on the x axis at time t the continuous Heisenberg group (with matrix multiplication) would implement constant velocity motion with velocity a together with a time-and-position translation.

After checking closure under matrix products and that inverses of matrices of this form are also of this form we see this is a group: another Lie group.

The "biggest" examples are the sets of invertible $n \times n$ matrices, the general linear groups $GL_n(\mathbb{R})$ or $GL_n(\mathbb{C})$.

The special linear groups $SL_n(\mathbb{R})$ or $SL_n(\mathbb{C})$, those matrices with determinant 1, are also examples.

The orthogonal matrices $\mathcal{O}_n(\mathbb{R})$ or $\mathcal{O}_n(\mathbb{C})$, consist of matrices M with $M^T M = I_n$. In the real case these are just the matrices whose columns form an orthonormal basis of $\mathbb{R}^n$.

The unitary matrices $\mathcal{U}_n$ consist of the complex matrices for which $M^* M = I_n$ where M^* is the conjugate transpose of M.

Recall dot product of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{C}^n$ is given by

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{n} \overline{x^i} y^i.$$

Columns of a unitary matrix form an orthonormal basis of $\mathbb{C}^n$.

Algebras

A Lie algebra of matrices is a real vector space V of matrices for which the Lie product of any two members of V is also in V.

$$[A, B] = AB - BA \in V \quad \text{whenever} \quad A, B \in V.$$

$\{0\}$ and $\mathbb{R}$ are zero and one dimensional Lie algebras, respectively.

$\mathbb{R}i$, the vector space of pure complex numbers, is also 1-dimensional.

$\mathbb{C}$ is a 2-dimensional Lie algebra.

$\mathbb{R}\vec{\boldsymbol{i}} \oplus \mathbb{R}\vec{\boldsymbol{j}} \oplus \mathbb{R}\vec{\boldsymbol{k}}$, the pure quaternions, are a 3-dimensional example. (It is easy, but one must check that if p and q are pure quaternions then $pq - qp$ has no real part.)

$\mathbb{H}$ itself is a 4-dimensional Lie algebra.

The 3-dimensional vector space of matrices of the form $\begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$ is a Lie algebra.

The biggest examples of Lie algebras are provided by $\mathbb{M}_{n\times n}(\mathbb{R})$ and $\mathbb{M}_{n\times n}(\mathbb{C})$.

Let $\mathcal{T}raceless_n(\mathbb{R})$ and $\mathcal{T}raceless_n(\mathbb{C})$ be the real and complex matrices with trace 0, respectively. Both are Lie algebras.[2]

Let $\mathcal{S}kew_n(\mathbb{R})$ and $\mathcal{S}kew_n(\mathbb{C})$ be the real and complex skew-symmetric matrices, respectively. Both are Lie algebras.

Let $\mathcal{S}kewHerm_n$ be the complex skew-Hermitian matrices, respectively. $M \in \mathcal{S}kewHerm_n$ exactly when $M^* = -M$. This is a Lie algebra.

[2]The Lie bracket of *any* two matrices has trace 0. The ijth entry of $XY - YX$ is $x_{i,k}y_{k,j} - y_{i,k}x_{k,j}$. The sum of the diagonal entries is then $x_{i,k}y_{k,i} - y_{i,k}x_{k,i} = 0$.

The Exponential and Logarithm

A series of matrices formed from a sequence of $m \times n$ matrices A_i is the sequence $S_k = \sum_{i=0}^{k} A_i$, called the sequence of partial sums of the A_i. The series is said to converge if the sequence of partial sums converges. When the series converges the limit matrix may be denoted $\sum_{i=0}^{\infty} A_i$. We are only interested in square matrices here.

We will use operator norm to describe convergence. This is comparable to the usual dot product "Frobenius" norm but more convenient for us because $\|I_n\| = 1$ and, like the Frobenius norm, $\|AB\| \leq \|A\|\|B\|$.

For square A we define e^A to be

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

This series has infinite radius of convergence and is continuous: if A is close to B then e^A is close to e^B.

Near the zero matrix the exponential is one-to-one.

The function $f(t) = e^{tA}$ is differentiable and the derivative can be calculated by term-by-term differentiation as Ae^{tA}.

If A and B commute then $e^{A+B} = e^A e^B$. This is *not* generally true if A and B fail to commute.

For square A we define $\ln(A)$ to be the limit of the series

$$\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j}(A - I_n)^j = -\sum_{j=1}^{\infty} \frac{1}{j}(I_n - A)^j$$

whenever that series converges. It *will* converge if $\|A - I_n\| < 1$.

$A = \ln(e^A)$ whenever $\left\|I_n - e^A\right\| < 1$ (Satisfied if $\|A\| < \ln(2)$).

$A = e^{\ln(A)}$ whenever $\|I_n - A\| < 1$.

There is an open neighborhood of the $\mathcal{N}_0$ of the zero matrix and an open neighborhood $\mathcal{N}_1$ of the identity matrix so that $e^{(\cdot)}$ and $\ln(\cdot)$ are inverse to each other when restricted to these neighborhoods.

The matrix differential equation $F' = MF$ has a unique matrix valued solution function $F(t) = e^{Mt}F(0)$ for each initial condition matrix $F(0)$.

$$e^{\operatorname{trace} A} = \det(e^A)$$

and so $\det(e^A) = 1$ if and only if $\operatorname{trace}(A) = 0$.

(i) e^A is invertible for any A.

(ii) e^A has positive determinant for any A.

(iii) If S is real and skew-symmetric then e^S is orthogonal.

(iv) If S is complex and skew-Hermitian then e^S is unitary.

(v) If $S \in \mathcal{T}raceless(\mathbb{F})$ then $e^S \in SL_n(\mathbb{F})$ where $\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$.

Lie Algebra ⇔ Lie Group

The Lie algebra **for** a Lie group is defined to be the set of all matrices that can be obtained as

$$\lim_{h \to 0} \frac{F(h) - F(0)}{h}$$

for differentiable function F entirely contained in the group and defined on some (possibly tiny) interval $(-\varepsilon, \varepsilon)$ around 0.

One proves that this set is a vector space and closed under Lie product. It then is proven that generic functions are not required: you can use functions of the form $F(t) = e^{tA}$ where A is the logarithm of a matrix in a neighborhood of 0 in the group. This map is entirely contained in the group (this is hard to show) for an interval of t values near 0. A basis of the Lie algebra can be selected from the set of these logarithms.

Exercise: The exponential maps a neighborhood of 0 in each Lie algebra listed earlier to a neighborhood of the identity in one (or more) of the Lie groups listed before. These Lie algebras are the Lie algebras **for** the associated Lie groups.

Find this pairing.

Note: the Pauli and gamma matrices generate Clifford algebras, which are also Lie algebras. Victor Polinger used the exponential of members of such algebras to create solutions to the Dirac equation for spin-one-half free particles and the Klein-Gordon equation for zero-spin free particles.