Our treatment of Lie groups will be concentrate mainly on Matrix Lie groups which will be defined later. But for now we simply give a taste of the general approach to Lie groups which views them as manifolds.

1 A brief Introduction to the general theory of Lie groups

Definition A **Lie group** is a differentiable manifold G which is also a group such that the group product

$$G \times G \to G$$

and the inverse map $g \to g^{-1}$ are differentiable.

This means that in order to understand the lie groups we will need to study manifolds.

Definition A manifold, $\mathcal{M}$ is a topological space such that for every point m in $\mathcal{M}$ there is a neighborhood U of m and a one to one continuous map ϕ to $\mathbb{R}^n$ i.e onto some open set $\phi(U)$ of $\mathbb{R}^n$ such that the inverse map $\phi^{-1} : \phi(U) \to U$ is also continuous. If the point m lies in intersection of two subsets of $\mathcal{M}$, U_i and U_j, with their respective maps ϕ_i and ϕ_j we demand that $\phi_i \circ \phi_j^{-1}$ is continuous and differentiable

The **tangent space** at m to $\mathcal{M}$, denoted as $T_m(\mathcal{M})$ is the set of all linear maps X from $C^\infty(\mathcal{M})$ into $\mathbb{R}$ satisfying:

1. "product rule": $X(fg) = X(f)g(m) + f(m)X(g) \quad \forall f$ and $g \in (C^\infty(\mathcal{M})$
2. "localization": If f is equal to g in a neighborhood of m, then $X(f) = X(g)$

An element of the tangent space is called the **tangent vector**. One can prove that if $x_1, \ldots x_n$ is a local coordinate system, then each tangent vector X at m can be expressed uniquely as

$$X(f) = \sum_{k=1}^{n} a_k \frac{\partial f}{\partial x_k}(m).$$

Thus is the manifold is n dimensional the tangent space is also n dimensional.

What we done is defined the notion of a tangent space without relying on the manifold being embedded in some larger euclidean space. The definition given above of the tangent vector is the generalization of the directional derivative. The reason we chose the directional derivative is that it does not depend on a curve moving through the specific point where it is defined.

1.1 Differentials of smooth mappings

A map Φ from a manifold $\mathcal{M}$ of dimension n to an manifold $\mathcal{N}$ of dimension m is called smooth if it is smooth for every local coordinates system ϕ_α on $\mathcal{M}$ and ϕ_β on $\mathcal{N}$ and the map $\phi_\beta \circ \Phi \circ \phi_\alpha^{-1}$ is a smooth map from an open subset of $\mathbb{R}^n$ to $\mathbb{R}^m$. Given a smooth map one can define a differential at every point m in $\mathcal{M}$, denoted by $\Phi_{*,m}$. This is the linear map of $T_m(\mathcal{M})$ into $T_{\Phi(m)}(\mathcal{M})$ given by

$$\Phi_{*,m}(X)(f) = X(f \circ \Phi)$$

where f is a smooth real valued function of $\mathcal{N}$, X is a tangent vector at m to $\mathcal{M}$. The map $\Phi_{*,m}$ will be a matrix of partial derivatives. An easy way to see this is that there are locally coordinates systems on $\mathcal{M}$ and $\mathcal{N}$, the map goes from $\mathcal{M}$ to $\mathcal{N}$ but X is defined with coordinates from M, so we need the jacobian.

Suppose that $\gamma : (a, b) \to \mathcal{M}$ is a smooth curve. Then for each $t \in (a, b)$ we will let $\frac{d\gamma}{dt}$ denote the element of $T_{\gamma(t)}(\mathcal{M})$ with the property that:

$$\frac{d\gamma}{dt}(f) = \frac{df(\gamma(t))}{dt} \tag{1}$$

for all smooth functions of $\mathcal{M}$. In a smooth local coordinate system $(x_1, \ldots, x_n)$, we can find smooth functions $x_1(t), \ldots, x_n(t)$ with $x_k(t)$ being $x_k(\gamma(t))$. The chain rule then says that $\frac{df(\gamma(t))}{dt} = \sum \frac{\partial f}{\partial x_k} \frac{dx_k}{dt}$. Combining this with the equation above we have that

$$\frac{d\gamma}{dt} = \sum \frac{\partial f}{\partial x_k} \frac{dx_k}{dt} \tag{2}$$

1.2 Vector fields

A **vector field** is a map X that associates to each point m in $\mathcal{M}$ a tangent vector $X_m \in T_m(\mathcal{M})$. In a local coordinate system we have:

$$X_m(f) = \sum_{k=1}^{n} a_k(m) \frac{\partial f}{\partial x_k} \tag{3}$$

where the a_k are real valued functions. We can apply a vector field X to a function f by applying X_m to f. The result is another function. So smooth vector field is a map from $C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ that satisfies the product rule in the form

$$X(fg) = fX(g) + X(f)g$$

in the above definition we are not evaluating f or g at any point. The equation can be summarized or restated as saying that a vector field is a **derivation** of the algebra of smooth functions.

We can view vector fields are first order differential operators, if we consider the product of two vector fields YX, this will be a second order differential operator but will not be a vector field but if we consider the commutator of the vector fields then the second order terms cancel out and we arrive again at a first order differential operator. So in other words the commutator produces another vector field. This is important fact when we arrive at describing the lie algebra of a lie group.

1.3 Flow Along a vector field

If X is a vector field and $\gamma : (a, b) \to \mathcal{M}$ is a smooth curve in $\mathcal{M}$, then γ is called an **integral curve** for X if for each $t \in (a, b)$, we have $\frac{d\gamma}{dt} = X_{\gamma(t)}$. In a smooth local coordinate system $x_1, x_2, \ldots x_n$, $\gamma(t)$ is represented by the family of functions $x_1(t), x_2(t), \ldots x_n(t)$. In light of (4.2) and (4.3) , $\frac{d\gamma}{dt} = X_{\gamma(t)}$ implies that

$$\frac{dx_k(t)}{dt} = a_k(x_1(t), \ldots, x_n(t)) \tag{4}$$

This is a system of ordinary differential equations(not necessarily linear) and applying results from differential equations we get existence and uniqueness.

A vector field X is called **complete** if $\gamma(t)$ can be defined for all t for initial points m. Any vector field on a compact manifold is always complete. If X is a complete vector fieldm then one can define the associated **flow** on $\mathcal{M}$. This sis a family of maps $\Phi_t : \mathcal{M} \to \mathcal{M}$ defined so that if γ is an integral curve for X with $\gamma(0) = m$, then $\Phi)_t(m) = \gamma(t)$. This means that $\Phi)_t(m)$ is defined starting at ma and "flowing" along the vector field X for time t. If X is a smooth complete vector field, then each Φ_t is a smooth map of $\mathcal{M}$ to itself, and he maps satisfy $\Phi_t \circ \Phi_s = \Phi_{t+s}$

We have already given the definition of a lie group and talked about lie algebras in some detail.

1.4 The Lie Algebra

If G is a Lie group and g an element of G, we define a map $L_g : G \to G$ by $L_g(h) = gh$, this is the "left multiplication by g" map. which is smooth since the product map $G \to G$ is assumed itself to be smooth. The differential of L_g denoted as $(L_g)_*$ will be a linear map of L_g at point h from $T_h(G)$ to $T_{gh}(G)$. A vector field is called **left invariant** if X satisfies:

$$(L_g)_*(X_h) = X_{gh} \tag{5}$$

Let $T_e(G)$ denote the tangent space at the identity then given any vector $v \in T_e(G)$ there is a unique left invariant vector field X^v with $X^v_e = v$ which can be constructed by defining

$$X^v_g = (L_g)_*(v) = (L_g)(X^v_e) \tag{6}$$

We can show that this vector field is left invariant thusly:

$$\begin{aligned}(L_h)_*(X_g^v) =& (L_h)_*(L_g)_*(v)\\ =& (L_{gh})_*(v)\\ =& (L_{gh})_*(X_e^v)\\ =& X_{gh}^v\end{aligned}$$

Definition The **Lie algebra** $\mathfrak{g}$ of a Lie group G is the tangent space at the identity with the bracket operation defined by

$$[v,w] = [X^v, X^w]_e$$

If we identitfy the space of left invariant vector fields with $T_e(G)$ by means of the map $v \longleftrightarrow X^v$, then $\mathfrak{g}$ is just the space of left invariant vector fields which forms a lie algebra under the commutator of vector fields.

1.5 Exponential Mapping

The exponential mapping for a general Lie group is defined in terms of the flow along left-invariant vector fields, this is justified because every left invariant vector field on G is complete.

Definition Let G be a lie group and let $\mathfrak{g} = T_e(G)$ be its lie algebra. For each $v \in \mathfrak{g}$, let X^v be the associated left-invariant vector field and let Φ_t^v be the associated flow. Then the exponential mapping is the map exp: $\mathfrak{g} \to G$ defined by

$$exp(v) = \Phi_t^v(e)$$

This means if we want to calculate exp(v) we construct the left invariant field X^v and we then calculate its integral curve γ^v that starts at the identity. Then $exp(v) = \gamma^v(1)$

1.5.1 Matrix Lie Groups as Lie Groups

To make the connection between Matrix Lie Groups and general lie groups, the integral curve we choose is $\gamma(t) = e^{tX}$ where X is an element in the lie algebra. This chosen so that it's value at the identity is X. To show that e^{tX} is an integral curve we need to show that

$$\frac{d\gamma(t)}{dt} = (L_{e^{tX}})_*(X) \tag{7}$$

Note: Remember for an integral cure $\gamma(t)$ we have that $\frac{d\gamma(t)}{dt} = X_{\gamma(t)}$

We proceed in the following manner:

$$\begin{aligned}\frac{d}{dt}e^{tX} =& \frac{d}{da}e^{(t+a)X}|_{a=0}\\ =&\frac{d}{da}e^{tX}e^{aX}|_{a=0}\\ =&\frac{d}{da}L_{e^{tX}}e^{aX}|_{a=0}\\ =&(L_{e^{tX}})_*\frac{d}{da}e^{aX}|_{a=0}\\ =&(L_{e^{tX}})_*(X)\end{aligned}$$

So every matrix Lie group is a Lie group.

2 Bridge between Lie Groups and Matrix Lie Groups

It is helpful to make a more concrete connection between looking at lie groups as manifolds and forgetting about their manifold structure.

We proceed by analyzing the affine group of the line A(1) with matrix representation

$$\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}$$

with $x > o$ or the right half of the plane.

Let $t \to h(t)$ be a curve of matrices in the group G with $h(0) = h$ and $h'(0) = X_h$. Since $G \subset GL(n)$, this curve is simply a matrix h whose entries $h_{jk}(t)$ are smooth functions of the parameter t. $h(t)$ describes a curve in an n^2 dimensional euclidean space.

X_h, the tangent to this curve is simply the matrix whose entries are the derivatives at $t = 0, h'_{jk}(0)$. h' does not need to in the group. For the constant matrix g, the curve $t \longmapsto gh(t)$ will have for tangent vector at $t = 0$ to the matrix.

$$L_{g_*}X_h = gh'(0) = gX_h \tag{8}$$

For every $\left(\begin{smallmatrix} x & y \\ 0 & 1 \end{smallmatrix}\right) \in A(1)$ we identify with $(x, y) \in \mathbb{R}^2$ and tangent vectors $\left(\begin{smallmatrix} \frac{dx}{dt} & \frac{dy}{dt} \\ 0 & 0 \end{smallmatrix}\right)$ we identify with $(\frac{dx}{dt}, \frac{dy}{dt})^T$

which is the tangent vector $(\frac{dx}{dt}\frac{\partial}{\partial x}, \frac{dy}{dt}\frac{\partial}{\partial y})$.

We us left translate the vectors $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ at the identity e to the point to (x, y)

Consider the curve $\left(\begin{smallmatrix} 1+t & 0 \\ 0 & \end{smallmatrix}\right)$ for $\frac{\partial}{\partial x}$ whose tangent at w is $\frac{\partial}{\partial x}$. Let g be the matrix $\left(\begin{smallmatrix} x & y \\ 0 & 1 \end{smallmatrix}\right) \in A(1)$ then we have

$$L_{g_*}\frac{\partial}{\partial x} = \frac{d}{dt}(gh(t))_{t=0} = \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} \tag{9}$$

The left translate of $\frac{\partial}{\partial x}$ to (x,y) is $X_1 = x\frac{\partial}{\partial x}$
Constructing the left translate of $\frac{\partial}{\partial y}$ at $(1,0)$ to the point (x,y)

$$L_{g_*}\frac{\partial}{\partial x} = \frac{d}{dt}\left(gh(t)\right)_{t=0} = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} \tag{10}$$

where now $h(t) = \left(\begin{smallmatrix} 0 & 1+t \\ 0 & 1 \end{smallmatrix}\right)$

We can consider the dual basis for the left invariant vector fields σ^1 and σ^2. The dual basis to $x\frac{\partial}{\partial x}$ and $x\frac{\partial}{\partial y}$ is $\sigma^1 = \frac{dx}{x}$ and $\sigma^2 = \frac{dy}{x}$. Thus for A(1) the left invariant area form or the left haar measure is

$$\sigma^1 \wedge \sigma^2 = \frac{dx \wedge dy}{x^2} \tag{11}$$

for any compact region U in A(1).

We also calculate the right translates in order to find the right invariant haar measure

$$R_{g_*}\frac{\partial}{\partial x} = \frac{d}{dt}\left(h(t)g\right)_{t=0} = \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix} \tag{12}$$

where is $h(t) = \left(\begin{smallmatrix} 1+t & 0 \\ 0 & 1 \end{smallmatrix}\right)$ So the right translate is $\tilde{X}_1 = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$. The same calculation with $h(t) = \left(\begin{smallmatrix} 0 & 1+t \\ 0 & 1 \end{smallmatrix}\right)$ gives $\frac{\partial}{\partial y}$. This gives the dual bases to be $\sigma^1 = \frac{1}{2}\left(\frac{1}{x}dx + \frac{1}{y}dy\right)$ and $\sigma^2 = dy$ and therefore the right invariant haar measure

$$\sigma^1 \wedge \sigma^2 = \frac{1}{2}\left(\frac{1}{x}dx \wedge dy\right) \tag{13}$$

2.1 One Parameter Subgroups

A one parameter subgroup of G is by definition a differential homomorphism (in particular, a path)

$$g : \mathbb{R} \to G$$

such that

$$t \longmapsto g(t) \in G$$

Or the additive group of the reals in to the group G. Thus $g(s+t) = g(s)g(t)$. So $g_{ij}(t+s) = \sum_k g_{ik}(t)g_{kj}(s)$. Differentiating both sides with respect to s and putting s=0 gives $g'(t) = g(t)g'(0)$ where $g'(0)$ is a constant matrix. The solution to this is $g(t) = g(0)exp\{tg'(0)\}$. If G is not a matrix group this is okay because the differential equation is really saying $g'(t) = L_{g(t)*}g'(0)$ i.e the tangent vector X to the 1-parameter subgroup is the left translated along the group. So given a tangent vector X_e at e in G, the *1-parameter subgroup* of G whose tangent at e is X_e is the integral curve through e of the vector field X on G resulting from left translation of X_e over all of G.

For $A(1)$ to find the 1-parameter subgroup having tangent vector $(a,b)^T$ at the identity, we left translate this vector over A(1). The left translate of $(a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y})$. So we solve

$$\begin{aligned}\frac{dx}{dt} &= ax \\ \frac{dy}{dt} &= bx \qquad \text{x(0)=1, y(0)=1}\end{aligned}$$

The solutions are $x(t) = e^{at}, y(t) = \frac{be^{at}-b}{a}$ so that we have

$$\begin{pmatrix} e^{at} & \frac{be^{at}-b}{a} \\ 0 & 1 \end{pmatrix} \tag{14}$$

3 Matrix Lie Groups

Definition Matrix Lie group is any subgroup G of $GL(n : \mathbb{C})$ with the following property: If A_m is any sequence of matrices in G, and A_m converges to some matrix A then either $A \in G$, or A is not invertible.

3.1 Matrix Exponential

We state without proof the following results

Theorem 3.1 *Let X be an $n \times n$ real or complex matrix X, the series $e^X = \sum_{m=0}^{m=\infty} \frac{X^m}{m!}$ converges.*

Proposition 3.2 *Let X and Y be n by n matrices, then the following are true:*
1.$e^0 = I$
2.$(e^X)^* = e^{X^*}$
3.e^X *is invertible and* $(e^X)^{-1} = e^{-X}$.
4.$e^{(\alpha+\beta)X} = e^{\alpha X}e^{\beta X}$ *for all* α *and* β
5. *If* $XY = YX$ *then* $e^{X+Y} = e^X e^Y$
6.*If* C*is invertile then* $e^{CXC^{-1}} = Ce^X C^{-1}$

Proof Using theorem 4.3.1, 1 and 2 should be obvious. Point 3 and point 4 follow from point 5. So we proceed to prove it.

$$\begin{aligned} e^{X+Y} &= \sum_{n=0} \frac{(X+Y)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{m=0}^{m=n} \binom{n}{m} X^m Y^{(n-m)} \\ &= \sum_{n=0} \frac{1}{n!} \sum_{m=0}^{m=n} \frac{n!}{m!(n-m)!} X^m Y^{(n-m)} \\ &= \sum_{n=0}^{n=\infty} \frac{1}{m!} X^m \sum_{m=0}^{m=n} \frac{1}{(n-m)!} Y^{(n-m)} \\ &= e^X e^Y \end{aligned}$$

Note the second step follows only because X and Y commute.
For point 6 we use theorem 4.3.1 on the left side.

$$
\begin{aligned}
e^{CXC^{-1}} =& I + CXC^{-1} + \frac{1}{2}(CXC^{-1})^2 + \frac{1}{3!}(CXC^{-1})^3 \dots \\
=& CC^{-1} + CXC^{-1} + \frac{1}{2}(CXC^{-1})^2 + \frac{1}{3!}(CXC^{-1})^3 \dots \\
=& C\left(I + X + \frac{1}{2}(X)^2 + \frac{1}{3!}(X)^3 \dots\right) C^{-1} \\
=& Ce^X C^{-1} \square
\end{aligned}
$$

For completeness sake we mention also that $\frac{d}{dt}e^{tX} = Xe^{tX}$.

3.1.1 Computing Matrix Exponential

It turns out there are three cases we need to consider when trying to compute the matrix exponential, these are if the matrix is diagonalizable, if it is nilpotent and lastly if it is arbitrary.

Case 1: Matrix is diagonalizable
Point 6 in the previous proposition offers a way of doing the calculation. We first diagonalize the matrix, then apply the result of point 6. The reason this helps is because computing the exponential of a diagonal matrix is highly trivial. First note that a square diagonal matrix multiplied by itself n times is equal to the diagonal entries raised to the n. Then applying the expansion of the series to the diagonal matrix means we just get in the matrix the exponential of the diagonal entries.

Case 2: Matrix is nilpotent A matrix M being nilpotent means that for some $n > 1, M^n = 0$. This means that the series expansion is not infinite and does in fact terminate.

Case 3: Matrix is arbtrary One use the SN decomposition first. It turns out that any matrix can be written as a sum of a diagonalizable matrix S and a nilpotent matrix N. Further more S and N commute. One can then use point 4 of the proposition to calculate the result.

A special case Sometimes it turns out that for some $n > 1, M^n = M$. In this case it is far easier not to diagonalize but deal with the series expansion directly. An example of this is calculate e^{iJ_y} where J_y is the angular momentum in the y direction of a spin 1 particle.

An important result often used in quantum physics and statistical mechanics is the following

Theorem 3.3 ***Lie product formula***. *Let X and Y be $n \times n$ matrices that in* *general do not commute. Then*

$$e^{X+Y} = \lim_{m\to\infty} \left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right)^m \tag{15}$$

Proof If we multiply the power series expansion of $e^{\frac{X}{m}} e^{\frac{y}{m}}$ and keep only terms linear in X and Y we get

$$e^{\frac{X}{m}} e^{\frac{Y}{m}} = I + \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)$$

For sufficiently large m in the domain of the logarithm we have

$$log\left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right) = log\left(I + \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right)$$

If m is large we can say that

$$log\left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right) = \frac{X}{m} + \frac{Y}{m} + O\left(||\frac{X}{m} + \frac{Y}{m} + \frac{1}{m^2}||^2\right)$$

Taking the exponential of both sides gives

$$\left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right) = exp\left(\frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right)$$
$$\left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right)^m = exp\left(X + Y + O\left(\frac{1}{m}\right)\right)$$

By continuity of the exponential, the result follows as m goes to infinity

3.2 The nature of some Lie Algberas as viewed from Lie groups

Lie algebras can of course be studied on their own and for their own sake as was done in the previous chapter. We now approach lie algebras from the angle of the lie groups.That is given a lie group what does its lie algebra look like.

3.2.1 Special Linear Group

It will be proved later on that $det(e^X) = e^{trX}$. This implies that if the determinant of $e^X$is 1 it must be that the trace of X is zero. Thus we arrive at the statement that the lie algebra for special linear group has trace 0.

3.2.2 Unitary group

Recall that elements in the unitary groups are matrices such that $U^{\dagger} = U^{-1}$. This means that $(e^X)^* = (e^X)^{-1} = e^{-X} = e^{X^*}$. We thus have that $X* = -X$ i.e the lie algebra for the unitary group is generated by anti-hermitian matrices. For the special unitary group we have that the lie algebra is spanned by matrices that have the additional property that they have trace being 0.

3.2.3 Orthogonal groups

For these groups we have that $O^T = O^{-1}$. Thus $(e^X)^T = e^{x^T} = (e^X)^{-1} = e^{-X}$. We therefore have the conclusion that the lie algebra for the orthogonal is spanned by matrices X, with $X^T = -X$. This condition forces the diagonal entries to be zero and thus have trace being zero.

3.2.4 Generalized Orthogonal groups

There are denoted as $O(n,k)$ and are $(n+k)\times(n+k)$ such that for a diagonal matrix d $A^T dA = d$ for $A \in O(n,k)$. D is a diagonal matrix with the first n entries consisting of 1 and the last being -1. Similar arguments can be made to show that the lie algebra consists of matrices such that $dX^T d = -X$ for X in the lie algebra.

3.2.5 General properties of the Lie Algebra

If X is in the lie algebra $\mathfrak{g}$ of the group G and A is in G then $Ae^{tX}A^{-1}$ is in $\mathfrak{g}$ for all t $\in \mathbb{R}$. This follows trivially from proposition 4.3.2. We can show that the commutator of two elements in the lie algebra is also in the lie algebra. This was an assumption when we looked at lie algebras. Consider the following

$$\frac{d}{dt}e^{tX}Ye^{-tX} = Xe^{tX}Ye^{-tX} - e^{tX}YXe^{-tX}$$

Now evaluate the above expression at t=0. To get XY - YX. Now $e^{tX}Ye^{-tX} \in \mathfrak{g}$ for all t. This can be deduced from the comments in the first paragraph of this subsection and we know from that the lie algebra is a real subspace of $M(\mathbb{C})$ and in particular it is a topological closed subset, it therefore follows that

$$XY - YX = \lim_{h\to 0}\frac{e^{hX}Ye^{-hX} - Y}{h}$$

belongs in $\mathfrak{g}$

Theorem 3.4 *Suppose we have a lie group homomorphism Φ from the groups G to H. We can show that there is lie algebra homomorphism ϕ attached with it.*

We sketch the proof as follows.

Proof Since Φ is a continuous lie group homomorphism, we know that $\Phi(e^{tX})$ is a one parameter subgroup for each X in the $\mathfrak{g}$.. So this means there is a unique element in H which write as e^{tZ}. We now claim there is a map ϕ such that $\phi(X) = Z$ and we show that is has all the properties of a lie algebra homomorphism.

1. By definition we have $\Phi(e^X) = e^{\phi(X)}$
2.Check that $\phi(tX) = t\phi(X)$

3. By 1 and 2 we have $e^{t\phi(X+Y)} = e^{\phi(t(X+Y))} = \Phi(e^{t(X+Y)})$
4.Use the Lie product formula to show that $\Phi(e^{t(X+Y)}) = e^{t(\phi(X)+\phi(Y)}$. The differentiate the result and evaluate at 0 to get $\phi(X+Y) = \phi(X) + \phi(Y)$
5.By 1 and 2 we have $e^{t\phi(AXA^-1)} = e^{\phi(tAXA^-1)} = \Phi(e^{tAXA^-1})$. Use the fact that Φ is a homomorphism to get $\Phi(e^{tAXA^-1}) = \Phi(A)e^{t\phi(X)}\Phi(A^{-1})$.
6. Recall that [X,Y] = $\frac{d}{dt}e^{tX}Ye^{-tX}|_{t=0}$. Apply the ϕ to both sides and use step 5 to get the result that $\phi([X,Y]) = [\phi(X), \phi(Y)]$ □

Carrying out all the steps in the sketch laid out gives the result we want. Now we note that the determinant is lie group homomorphism and the trace is a lie algebra homomorphism. We thus have the result that $det(e^X) = e^{trX}$.

3.2.6 The Adjoint map

Definition (The adjoint map). Let G be a matrix Lie group, with lie algebra $\mathfrak{g}$. Then for each $A \in G$ we define the linear map $Ad_A : \mathfrak{g} \to \mathfrak{g}$ such that $Ad_A(X) = AXA^{-1}$

Theorem 3.5 *Let G be a matrix Lie Group with lie algebra* $\mathfrak{g}$. *Let* $GL(\mathfrak{g})$ *denote the group of all invertible linear transformation of* $\mathfrak{g}$. *The for each* $A \in G, Ad_A$ *is an invertible linear transformation of* $\mathfrak{g}$ *with inverse* $Ad_{A^{-1}}$ *and the map* $A \to Ad_A$ *is a group homomorphism of G into* $GL(\mathfrak{g})$ *and satisfies* $Ad_A([X,Y]) = [Ad_A(X), Ad_A(Y)]$

Proof It should be easy to see that $Ad_{A^{-1}}$ is the inverse of Ad_A. It should be easy to check that $Ad_{AB} = Ad_A Ad_B$.

$$\begin{aligned}
Ad_A([X,Y]) =& A(XY)A^{-1} - A(YX)A^{-1} \\
=& A(XA^{-1}AY)A^{-1} - A(YA^{-1}AX)A^{-1} \\
=& (Ad_A([X,Y]) \\
=& [Ad_A(X), Ad_A(Y)]
\end{aligned}$$

By theorem 4.3.4 we know there is a lie algebra homomorphism from the lie algebra to the linear algebra of linear transformation acting on the lie algebra, which we denote as ad and defined such that $ad : \mathfrak{g} \to gl(\mathfrak{g})$. This was defined earlier in the context of lie algebras but we now show it follows from the lie group homomorphism. We can see this by considering the following $ad_X = \frac{d}{dt}Ad_{e^{tX}}|_{t=0}(Y) = \frac{d}{dt}e^{tX}Ye^{-tX}|_{t=0}$. But we have already carried out this calculation we know the result is [X,Y].

From this we can prove a highly no trivial result. Consider $e^{ad_X}(Y) = Ad_{e^X}(Y) = e^XYe^{-X}$. So we have that $(e^{ad_X}(Y) = e^XYe^{-X}$

4 Baker-Campbell Hausdorff Formula

The main purpose of the Baker-Campbell Hausdorff formula(BCH) is it shows that if ϕ is a lie algebra homomorphism from lie algebras $\mathfrak{g} \to \mathfrak{h}$ then if there is a map between groups G and H $\Phi : G \to H$ with lie algebras $\mathfrak{g} \to \mathfrak{h}$ respectively with $\Phi(e^X) = e^{\phi(X)}$ then Φ is a lie group homomorphism. It accomplishes this task by expressing the product of two lie group elements in terms of the lie bracket of the lie algebra.

The Baker-Campbell Hausdorff formula says

$$log(e^X e^Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{2}[Y,[X,Y]] + \dots \tag{16}$$

Now apply the lie algebra homomorphism ϕ to 4.9 we have

$$\phi(log(e^X e^Y)) = \phi(X) + \phi(Y) + \frac{1}{2}[\phi(X),\phi(Y)] + \frac{1}{12}[\phi(X),[\phi(X),\phi(Y)]] -$$

$$\frac{1}{2}[\phi(Y),[\phi(X),\phi(Y)]] + \dots$$

$$= log(e^{\phi(X)} e^{\phi(Y)})$$

Since we have that $e^X e^Y = e^{log(e^X e^Y)}$ then we have $\Phi(e^X e^Y) = e^{\phi}(log(e^X e^Y))$ then applying (BCH) we have that $\Phi(e^X e^Y) = \Phi(e^X)\Phi(e^Y)$, thus making Φ a lie group homomorphism.

Note: BCH holds for sufficiently small X and Y.

Let $g(z) = \frac{logz}{1-\frac{1}{z}}$ for some complex number z. This function is defined and analytic in the disk $|Z - 1| < 1$ and can be expressed as the laurent series $g(z) = \sum_{m=0}^{m=\infty} a_m(z-1)^m$. For a finite dimensional vector space V and an operator A on V the for $||A - I|| < 1$ we can define $g(A) = \sum_{m=0}^{m=\infty} a_m(A - I)^m$. There is an integral version of BCH and it is the following

$$log(e^X e^Y) = X + \int_0^1 g(e^{ad_X} e^{tad_Y})(Y)dt \tag{17}$$

Again the result holds for sufficiently small $||X||, ||Y||$ and X, Y $n \times n$ matrices.

4.1 Derivative of the Exponential Map

Before we deal with BCH we look at the derivative of the exponential map. Note that if X and Y commute then $e^{X+tY} = e^X e^{tY}$ and calculating $\frac{d}{dt}(e^{X+tY}|_{t=0})$ is trivial to calculate. But if they do not commute then the derivative is highly non-trivial to calculate.

Theorem 4.1 *Let X and Y be $n \times n$ complex matrices. Then,*

$$\frac{d}{dt} e^{X+tY}|_{t=0} = e^X \{ \frac{I - e^{-ad_X}}{ad_X}(Y) \} \tag{18}$$

and more generally

$$\frac{d}{dt} e^{X(t)} = e^{X(t)} \{ \frac{I - e^{-ad_{x(t)}}}{ad_{X(t)}}(\frac{dX}{dt}) \} \tag{19}$$

Proof Define $\Delta(X,Y) = \frac{d}{dt}e^{X+tY}|_{t=0}$ and note that

$$e^{X+tY} = \left(e^{\frac{X}{m}+\frac{tY}{m}}\right)^m \tag{20}$$

Now if differentiate 4.20 with respect to t and evaluate the expression at t=0. We will get m terms and in each term there will be one factor that has the derivative and the others will not. Once we evaluate the expressions at t=0 a generic terms will look like this

$$(e^{X/m})^n \left(\frac{d}{dt}e^{X/m+tY/m}|_{t=0}\right)(e^{X/m})^k \tag{21}$$

where $n+k = m-1$

Thus we have

$$\begin{aligned}\frac{d}{dt}e^{X+tY}|_{t=0} &= \sum_{k=0}^{m-1}\left(e^{\frac{X}{m}}\right)^{m-k-1}\left(\frac{d}{dt}e^{X/m+tY/m}|_{t=0}\right)\left(e^{\frac{X}{m}}\right)^k \\ &= e^{(m-1)/mX}\sum_{k=0}^{m-1} e^{(\frac{-X}{m})k}\Delta(X/m, Y/m)e^{(\frac{X}{m})k} \\ &= e^{(m-1)/mX}\frac{1}{m}\sum_{k=0}^{m-1}\left(e^{\frac{-ad}{m}}\right)^k(\Delta(X/m, Y))\end{aligned}$$

Linearity of $\Delta(X,Y)$ was used in the last step. The above expression is true for all m and in particular true as $m \to \infty$, we have that $\Delta(0,Y) = Y$ In the limit as m goes to infinity we are left with studying

$$\frac{1}{m}\sum_{k=0}^{m-1}\left(e^{-\frac{-ad_X}{m}}\right)^k \tag{22}$$

We proceed with a non-rigorous argument by pretending the summand is a number rather than an operator and apply the geometric series formula. This is ultimately okay because we are dealing with operators with small enough norms. So we have

$$\frac{1}{m}\sum_{k=0}^{m-1}\left(e^{-\frac{-ad_X}{m}}\right)^k = \frac{1}{m}\frac{1-e^{-ad_X}}{1-e^{-ad_X/m}} \tag{23}$$

and as m goes to infinity we have $\frac{1-e^{-ad_X}}{ad_X}$ □

We now turn to the proof the Baker Campbell Hausdorff formula.

4.2 Proof of Baker Campbell Hausdorff Formula

Define $Z(t) = log(e^X e^{tY})$ with t between 0 and 1. We need to compute $Z(1)$. We have that

$$e^{-Z(t)}\frac{d}{dt}e^{Z(t)} = \left(e^X e^{tY}\right)^{-1}e^X e^{tY}Y = Y$$

One the other hand we have just shown that

$$e^{-Z(t)}\frac{d}{dt}e^{Z(t)} = \frac{I-e^{-ad_{Z(t)}}}{ad_{Z(t)}}\left(\frac{dZ}{dt}\right)$$

So we have

$$\frac{I-e^{-ad_{Z(t)}}}{ad_{Z(t)}}\left(\frac{dZ}{dt}\right) = Y$$

X and Y are small so we can invert the formula to get

$$\frac{dZ}{dt} = \left(\frac{I-e^{-ad_{Z(t)}}}{ad_{Z(t)}}\right)^{-1}(Y) \tag{24}$$

By the Ad homomorphism and the relation between "Ad" and "ad" we have $e^{ad_{Z(t)}} = e^{ad_X}e^{ad_Y}$ or $ad_{Z(t)} = log\left(e^{ad_X}e^{ad_Y}\right)$. Plugging these results into 4.17 we have

$$\frac{dZ}{dt} = \left(\frac{I-e^{ad_X}e^{ad_Y)}}{log\left(e^{ad_X}e^{ad_Y}\right)}\right)^{-1}(Y) \tag{25}$$

Remember before we defined the function $g(z) = \left(\frac{1-z^{-1}}{logz}\right)^{-1}$. Applying this to 4.18 we have

$$\frac{dZ}{dt} = g(e^{ad_X}e^{ad_Y})(Y) \tag{26}$$

Note that Z(0)=X and integrating we get

$$Z(1) = X + \int_0^1 g(e^{ad_X}e^{tad_Y})(Y)dt \tag{27}$$

which is what we want.

We can make connection with the series representation of BCH in the following manner:
The closed form of the expression for the series expansion of g is

$$g(z) = 1 + \sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{m(m+1)}(z-1)^m$$

we first deal with the term (Z-1)

$$e^{ad_X}e^{tad_Y} - I = ad_X + tad_Y + tad_x ad_Y + \frac{(ad_X)^2}{2} + t^2\frac{(ad_Y)^2}{2} + t^2\frac{(ad_Y)^2}{2} + \ldots \tag{28}$$

so we have

$$
\begin{aligned}
g(e^{ad_X}e^{tad_Y}) =& I + \frac{1}{2}\left(ad_X + tad_Y + tad_X ad_Y + \frac{(ad_X)^2}{2} + t^2\frac{(ad_Y)^2}{2} + t^2\frac{(ad_Y)^2}{2} + \dots\right) \\
&- \frac{1}{6}\left(ad_X + tad_Y + \dots\right)^2 + \dots
\end{aligned}
$$

We then have (neglecting high order terms)

$$
\begin{aligned}
log(e^X e^Y) =& X + \int_0^1 \left[Y + \frac{1}{2}[X,Y] + \frac{1}{4}[X,[X,Y]] - \frac{1}{6}[X,[X,Y]] - \frac{t}{6}[Y,[X,Y]]\right] dt \qquad (29) \\
=& X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[]X,Y] + \text{ high order terms} \qquad (30)
\end{aligned}
$$