LIE GROUPS

SAI CHINTAGUNTA AND YOANN IAIROS

Contents

1. Introduction 1
2. Preliminary Notions 1
3. Matrix Groups 5
4. The Exponential Function 5
5. Lie Algebras 10
6. Smooth Manifolds and Tangent Spaces 14
7. Lie Groups 18
8. The Rogers-Ramanujan Identities 20
Acknowledgements 21
References 21

1. Introduction

In Lie Theory, we would like to prove things concerning groups using tools from analysis. The exponential function, which is of great importance in analysis, can be generalized so as to make sense for matrices and it provides the link between lie algebras and matrix groups (every matrix group is a Lie group, and although the converse is false, many interesting results can be arrived at just by considering matrix groups). Lie theory is a subject where the study of examples is very important. Accordingly, we shall give examples when the occasions to do so arise.

After putting metrics on our matrices in Section 2, we define matrix groups in Section 3. We show in Section 4 how the exponential function is a locally invertible map that sends matrices to invertible matrices, and in Section 5 we demonstrate how the exp function provides the link between matrix groups and their Lie algebras. In Section 6, we explore smooth manifolds, their tangent spaces, and derivatives of smooth maps between them. Finally, in Section 7 we define Lie groups and explore their connections to Lie algebras and matrix groups.

2. Preliminary Notions

Throughout the entire paper, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. We will also freely use the term 'small' (respectively 'large') to mean 'sufficiently small' ('sufficiently large') whenever needed. We will denote the zero $n \times n$ matrix by O. Also, the entry of a matrix A situated at the intersection of the i-th row and j-th column will be denoted by A_{ij}.

Let $M_n(\mathbb{K})$ denote the set of all $n \times n$ matrices over $\mathbb{K}$. Then, matrix addition gives an abelian group structure on $M_n(\mathbb{K})$ and $M_n(\mathbb{K})$ becomes an n^2-dimensional vector space over

Date: August 17, 2025.

$\mathbb{K}$ with scalar multiplication. Now a norm on a vector space naturally induces a metric which, in turn, induces a topology. Turning our vector space into a topological space enables us to do analysis on it. Therefore we are going to turn $M_n(\mathbb{K})$ into a normed vector space. Now it turns out that the metric topology we wish to induce on $M_n(\mathbb{K})$, which is a vector space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ of finite dimension n^2, does not depend on our choice of norm (i.e. the topologies induced are the same, in the sense that the open subsets of $M_n(\mathbb{K})$ are the same in each instance), thanks to the following theorem:

Theorem 2.1. *If we have two norms $\|\cdot\|$ and $\|\cdot\|'$ over a finite dimensional vector space V over $\mathbb{R}$ or $\mathbb{C}$, then they define equivalent topologies. In particular, there exist constants $A, B > 0$ such that $A\|v\|' \leq \|v\| \leq B\|v\|'$ for all $v \in V$.*

For the proof the reader is referred to [Con25].

Remark 2.2. By the last theorem we do not need to specify the topology when speaking about convergence of sequences in $M_n(\mathbb{K})$.

To define a convenient norm on $M_n(\mathbb{K})$, we first define the usual norm $|\cdot|$ on $\mathbb{K}$ as follows: for $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$, we define $|\vec{x}| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$. We now use this norm on $\mathbb{K}^n$ to define a norm $\|\cdot\|$ on $M_n(\mathbb{K})$ as follows: given $A \in M_n(\mathbb{K})$, we define $\|A\| = \max \mathcal{S}_A$, called the *operator norm*, where $\mathcal{S}_A = \{\frac{|A\vec{x}|}{|\vec{x}|} : \vec{0} \neq \vec{x} \in \mathbb{K}^n\} = \{|A\vec{x}| : \vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\}$. Of course, we must now that that this is well-defined and that it is indeed a norm. Since $\{\vec{x} \in \mathbb{K}^n : |\vec{x}| = 1\}$ is a compact subset of $\mathbb{K}^n$, therefore its image under the linear transformation A is also compact, and therefore $S_A = \{|A\vec{x}| : |\vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\}$ is a compact and so it has a maximum value, showing that $\|A\|$ is indeed well-defined. Next we show that it is indeed a norm.

Proposition 2.3. *The operator norm is a norm.*

Proof. Since $\|A\| = \max\{|A\vec{x}| : \vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\}$ is the maximum of a set of nonnegative numbers, we have $\|A\| \geq 0$. Secondly, the maximum is 0 if and only if $|A\vec{x}| = 0$ for all $|\vec{x}|$ which occurs if and only if $A = 0$. Next let $\lambda \in \mathbb{K}$. Then $\|\lambda A\| = \max\{|\lambda A\vec{x}| : \vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\} = |\lambda| \max\{|A\vec{x}| : \vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\} = |\lambda|\|A\|$. Finally, $\|A + B\| = \max\{|(A+B)\vec{x}| : \vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\} \leq \max\{|A\vec{x}| + |B\vec{x}| : \vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\} \leq \max\{|A\vec{x}| : |\vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\} + \max\{|B\vec{x}| : |\vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\} = \|A\| + \|B\|$. ■

We prove one more property of the operator norm $\|\cdot\|$ before moving on.

Proposition 2.4. *For $A, B \in M_n(\mathbb{K})$, we have $\|AB\| \leq \|A\|\|B\|$.*

Proof. Since $A = \max\{\frac{|A\vec{x}|}{|\vec{x}|} : \vec{0} \neq \vec{x} \in \mathbb{K}^n\}$, we have that $|A\vec{x}| \leq |\vec{x}|\|A\|$ for all $\vec{0} \neq \vec{x} \in \mathbb{K}^n$. Therefore $|AB\vec{x}| \leq |B\vec{x}|\|A\|$. Also $|B\vec{x}| \leq |\vec{x}|\|B\|$. Therefore $|AB\vec{x}| \leq \|A\|\|B\||\vec{x}|$, and in particular $\|AB\| \leq \|A\|\|B\|$, as required. ■

Therefore, we have successfully turned $M_n(\mathbb{K})$ into a normed vector space. Hence, it is now a metric space under the naturally induced metric $\rho : M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to \mathbb{R}_{\geq 0}$ defined by $\rho(A, B) = \|A - B\|$ for every $A, B \in M_n(\mathbb{K})$. This metric now naturally induces a topology on $M_n(\mathbb{K})$. We can therefore now speak of open and closed sets of $M_n(\mathbb{K})$, convergence of sequences in $M_n(\mathbb{K})$, as well as continuous functions $f : Y \subseteq M_n(\mathbb{K}) \to X$, where X

is a topological space and the subset $Y \subseteq M_n(\mathbb{K})$ has the subspace topology, in the usual manner. We will now give certain important examples of continuous functions.

Proposition 2.5. *The coordinate function* $\mathrm{coord}_{rs} : M_n(\mathbb{K}) \to \mathbb{K}$ *defined by* $\mathrm{coord}_{rs}(A) = A_{rs}$ *is a continuous function.*

Proof. To simplify notation, we consider the case $n = 2$ only, as other cases are exactly analogous, and we also consider only coord_{11}. We must show that given any $\epsilon > 0$, there exists $\delta > 0$ such that $|A'_{11} - A| < \epsilon$ whenever $\|A' - A\| < \delta$. We claim that we can take $\epsilon = \delta$. This is because $|A_{11}| \leq \sqrt{|A_{11}|^2 + |A_{12}|^2} = \left|\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}\right| = \left|A\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right| \leq \|A\|$, and so $|A'_{11} - A_{11}| \leq \|A' - A\|$. ■

As a first corollary of the fact that $|A_{ij}| \leq \|A\|$, it follows that:

Corollary 2.6. *If* $f : \mathbb{K}^{n^2} \to \mathbb{K}$ *is a continuous function, then the function* $F : M_n(\mathbb{K}) \to \mathbb{K}$ *defined by* $F(A) = f((A_{ij})_{1 \leq i,j \leq n})$ *is continuous.*

Next, since polynomials $f : \mathbb{C}^n \to \mathbb{C}$ over $\mathbb{C}$ are continuous functions, it follows that:

Corollary 2.7. *The determinant* $\det : M_n(\mathbb{K}) \to \mathbb{K}$ *and the trace* $\mathrm{Tr} : M_n(\mathbb{K}) \to \mathbb{K}$ *are continuous functions.*

So far we have proved that $|A_{ij}| \leq \|A\|$. But we can also give an upper bound on $\|A\|$.

Proposition 2.8. *For* $A \in M_n(\mathbb{K})$*, we have*

$$\|A\| \leq \sum_{i,j=1}^{n} |A_{ij}|.$$

Proof. Again we let $n = 2$ for simplifying notation but the general case is exactly analogous. Let $\vec{x} \in \mathbb{K}^2$ be such that $|\vec{x}| = 1$ and let $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$. Then

$$\begin{aligned} A\vec{x} &= \left|x_1 A\begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 A\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right| \leq \left|x_1 A\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right| + \left|x_2 A\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right| \leq \left|A\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right| + \left|A\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right| \\ &= \left|\begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix}\right| + \left|\begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix}\right| = \sqrt{|A_{11}|^2 + |A_{12}|^2} + \sqrt{|A_{21}|^2 + |A_{22}|^2} \\ &\leq |A_{11}| + |A_{12}| + |A_{21}| + |A_{22}| \end{aligned}$$

Since $\|A\| = \max\{|A\vec{x}| : \vec{x} \in \mathbb{K}^n, |\vec{x}| = 1\}$, it follows that $\|A\| \leq |A_{11}| + |A_{12}| + |A_{21}| + |A_{22}|$. ■

Next we prove that $M_n(\mathbb{K})$ is a complete space with respect to the norm $\|\cdot\|$.

Theorem 2.9. *Every Cauchy sequence* $\{A_r\}_{r \geq 0}$ *in* $M_n(\mathbb{K})$ *has a unique limit* $\lim_{r\to\infty} A_r$ *in* $M_n(\mathbb{K})$*. Furthermore,* $(\lim_{r\to\infty} A_r)_{ij} = \lim_{r\to\infty}(A_r)_{ij}$*.*

Proof. In metric spaces, convergent sequences converge to unique limits. We have already shown that $|A_{ij}| \leq \|A\|$ and so it follows that $\{(A_r)_{ij}\}_{r \geq 0}$ is a Cauchy sequence of complex numbers and so $\lim_{r\to\infty}(A_r)_{ij}$ is well-defined. What we remain to show therefore is that $\{A_r\}_{r \geq 0}$ converges to the matrix A for which $A_{ij} = \lim_{r\to\infty}(A_r)_{ij}$. For this, we consider the sequence $\{A_r - A\}_{r \geq 0}$. It follows from Proposition 2.8 that, as $r \to \infty$, we have $\|A_r - A\| \leq \sum_{i,j=1}^{n} |(A_r)_{ij} - A_{ij}| \to 0$ and so $A_r \to A$ as required. ■

We next define two important groups.

Definition 2.10. The *general linear group* $\mathrm{GL}_n(\mathbb{K}) \subseteq M_n(\mathbb{K})$ is defined as the set $\{A \in M_n(\mathbb{K}) : \det A \neq 0\}$. The special linear group $\mathrm{SL}_n(\mathbb{K}) \subseteq \mathrm{GL}_n(\mathbb{K})$ is the set $\{A \in M_n(\mathbb{K}) : \det A = 1\}$.

The reader will have no difficulty ascertaining that these are indeed groups with respect to the usual matrix multiplication.

Proposition 2.11. $\mathrm{GL}_n(\mathbb{K}) \subseteq M_n(\mathbb{K})$ *is an open subset and* $\mathrm{SL}_n(\mathbb{K}) \subseteq M_n(\mathbb{K})$ *is a closed subset.*

Proof. We have already seen in Corollary 2.7 that the determinant is a continuous function. we have $\mathrm{GL}_n(\mathbb{K}) = M_n(\mathbb{K}) \setminus \det^{-1}\{0\}$. But $\{0\}$ is a closed subset of $\mathbb{K}$. Hence $\det^{-1}\{0\}$ is closed and therefore $\mathrm{GL}_n(\mathbb{K})$ is open. Similarly $\mathrm{SL}_n(\mathbb{K}) = \det^{-1}\{1\}$ is closed in $M_n(\mathbb{K})$. Also since the complement of $\mathrm{SL}_n(\mathbb{K})$ in $\mathrm{GL}_n(\mathbb{K})$ is given by $(M_n(\mathbb{K}) \setminus \det^{-1}\{1\}) \cap \mathrm{GL}_n(\mathbb{K})$ and $M_n(\mathbb{K}) \setminus \det^{-1}\{1\}$ is open in $M_n(\mathbb{K})$, it follows that $\mathrm{SL}_n(\mathbb{K})$ is closed in $\mathrm{GL}_n(\mathbb{K})$. ■

We next define a new norm on $M_n(\mathbb{K})$.

Definition 2.12 (Hilbert-Schmidt norm)**.** For $A \in M_{\mathbb{K}}$ we define the norm

$$\|A\|_{HS} = \sqrt{\sum_{i \leq i,j \leq n} |A_{ij}|^2}.$$

Remark 2.13. $\|A\|_{HS}$ is just the usual norm on $\mathbb{K}^{n^2}$. It follows, in view of Theorem 2.1, that continuity of a function $f : M_m(\mathbb{K}) \to M_n(\mathbb{K})$ can be regarded as continuity in $f : \mathbb{K}^{m^2} \to \mathbb{K}^{n^2}$. In particular, f is continuous if and only if each of its component functions is continuous.

We next give some terminology that will be used to define topological groups shortly.

Definition 2.14. We defined the addition $\mathrm{add} : M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to M_n(\mathbb{K})$, multiplication $\mathrm{mult} : M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to M_n(\mathbb{K})$ and inversion $\mathrm{inv} : \mathrm{GL}_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K})$ maps as: $\mathrm{add}(X, Y) = X + Y, \mathrm{mult}(X, Y) = XY$ and $\mathrm{inv}(A) = A^{-1}$.

We can put the product topology as usual on $M_n(\mathbb{K}) \times M_n(\mathbb{K})$ by defining the metric $d((X_1, Y_1), (X_2, Y_2)) = \|X_1 - X_2\| + \|Y_1 - Y_2\|$ on it. We can now prove the following:

Proposition 2.15. *The maps* $\mathrm{add}, \mathrm{mult}$ *and* inv *are continuous.*

Proof. Continuity for add and mult follows because the entries of the output are polynomial functions of the entries of the inputs. Similarly, the inverse map is continuous since each entry of A^{-1} has the form (polynomial in the entries A_{ij})$\setminus \det A$ and since this is a continuous function of the entries of A it is a continuous function of A. ■

We now define a topological group.

Definition 2.16 (Topological Group)**.** Let G be a topological space and view $G \times G$ as the product space. Suppose that G is also a group with multiplication map $\mathrm{mult} : G \times G \to G$ and inverse map $\mathrm{inv} : G \to G$. Then G is a *topological group* if mult and inv are continuous.

For example, it trivially follows that any group G equipped with the discrete topology is a topological group. By what has been shown above, the following theorem automatically follows:

Theorem 2.17. *Each of the groups* $\mathrm{GL}_n(\mathbb{K})$ *and* $\mathrm{SL}_n(\mathbb{K})$ *is a topological group.*

3. Matrix Groups

We are now in a position to define matrix groups.

Definition 3.1 (Matrix Group)**.** A subgroup $G \subseteq \mathrm{GL}_n(\mathbb{K})$ which is also a closed subspace is a *matrix group over* $\mathbb{K}$ or a $\mathbb{K}$-*matrix group*. In order to emphasize the value of n , we will sometimes say that G is a *matrix subgroup* of $\mathrm{GL}_n(\mathbb{K})$.

We have the following proposition.

Proposition 3.2. *Let $G \subseteq \mathrm{GL}_n(\mathbb{K})$ be a matrix subgroup. Then a closed subgroup $H \subseteq G$ is a matrix subgroup H of* $\mathrm{GL}_n(\mathbb{K})$.

Proof. Since we have a metric topology on $M_n(\mathbb{K})$, we only have to show that H is closed in $\mathrm{GL}_N(\mathbb{K})$. Every sequence $\{A_n\}_{n\geq 0}$ in H with a limit in $\mathrm{GL}_n(\mathbb{K})$ actually has its limit in G since $A_n \in H \subseteq G$ for every n and G is closed in $\mathrm{GL}_n(\mathbb{K})$ by definition of a matrix group. Since H is closed in G by hypothesis, this means that $\{A_n\}_{n\geq 0}$ has a imit in H. So H is closed in $\mathrm{GL}_n(\mathbb{K})$, as desired. ■

We now give another definition.

Definition 3.3. A closed subgroup $H \subseteq G$ of a matrix group G is called a *matrix subgroup of* G.

Then a straightforward generalization of the last definition is:

Proposition 3.4. *Let G be a matrix group and let $K \subseteq G$ be a matrix subgroup of G. Then if $H \subseteq K$ is a matrix subgroup of K, H is also a matrix subgroup of G.*

Example. $\mathrm{SL}_n(\mathbb{K})$ is a matrix group. This is because we have already shown previously that $\mathrm{SL}_n(\mathbb{K}) \subseteq \mathrm{GL}_n(\mathbb{K})$ is closed in $\mathrm{GL}_n(\mathbb{K})$.

Example. The orthogonal group $O(n) = \{A \in \mathrm{GL}_n(\mathbb{K}) : A^T A = I\}$ is a matrix subgroup of $\mathrm{GL}_n(\mathbb{K})$.

4. The Exponential Function

The exponential function is very important and will enable us to link a matrix group with its Lie algebra (a term which will be defined later).

Definition 4.1. The exponential function on matrices are defined as follows:

$$\exp(A) = \sum_{n\geq 0} \frac{1}{n!} A^n = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots .$$

We will prove that the series for $\exp(A)$ converges for all $A \in M_n(\mathbb{K})$. We note first that $\sum_{n=1}^{\infty} A_n$ converges if its partial sums converge (that is the definition of convergence we are employing). We also say that it is *absolutely convergent* if $\sum_{n=1}^{\infty} \|A_n\|$ converges in the usual way. Now we have the following proposition:

Proposition 4.2. *Every absolutely convergent series in $M_n(\mathbb{K})$ is convergent.*

Proof. Suppose that $\sum_{n=1}^{\infty} \|A_n\|$ is convergent. Then the sequence of partial sums is a Cauchy sequence and so, given any $\epsilon > 0$, we can find some N for which $\sum_{k=N+1}^{\infty} \|A_k\| < \epsilon$. Now let $S_m = \sum_{k=1}^{m} A_k$ be the m-th partial sum. Then whenever $m > n > N$ we have

$$\|S_m - S_n\| = \left\| \sum_{k=1}^{m} A_m - \sum_{k=1}^{n} A_k \right\| = \left\| \sum_{k=n+1}^{m} A_k \right\| \leq \sum_{n+1}^{m} \|A_k\| \leq \sum_{k=N+1}^{\infty} \|A_k\| < \epsilon.$$

The partial sums S_m therefore form a Cauchy sequence and therefore the series converges by Theorem 2.9. ■

Now we prove that $\exp(A)$ converges for all $A \in M_n(\mathbb{K})$.

Proposition 4.3. *The series for* $\exp(A)$ *converges for all* $A \in M_n(\mathbb{K})$ *and furthermore* $\exp(A)$ *is a continuous function.*

Proof. From Proposition 2.4, it follows that $\|A^n\| \leq \|A\|^n$ for nonnegative integers n. Hence,

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

which is absolutely convergent. Hence, by Proposition 4.2, the infinite series for $\exp(A)$ is convergent. Continuity follows by uniform convergence on bounded sets of matrices and the fact that the partial sums are continuous. ■

We have the following proposition:

Proposition 4.4. (1) *If* $A, B \in M_n(\mathbb{K})$ *commute then* $\exp(A+B) = \exp(A)\exp(B)$.
(2) $\exp(A) \in \mathrm{GL}_n(\mathbb{K})$ *and* $\exp(A)^{-1} = \exp(-A)$.

Proof. (1) The absolute convergence of exp enables us to make use of the Cauchy product formula. We have

$$\begin{aligned}
\exp(A)\exp(B) &= \sum_{m\geq 0} \frac{1}{m!} A^m \sum_{n\geq 0} \frac{1}{n!} B^n \\
&= \sum_{m} \sum_{n} \frac{1}{m!n!} A^m B^n \\
&= \sum_{k\geq 0} \sum_{m=0}^{k} \frac{1}{m!(k-m)!} A^m B^{k-m} \\
&= \sum_{k\geq 0} \frac{1}{k!} \sum_{m=0}^{k} \frac{k!}{m!(k-m)!} A^m B^{k-m} \\
&= \sum_{k\geq 0} \frac{1}{k!} (A+B)^k \\
&= \exp(A+B),
\end{aligned}$$

where the use of the Binomial Theorem in the penultimate step has crucially relied on the assumption that A and B commute.

(2) Since A and $-A$ commute, the previous part implies that

$$\exp(A)\exp(-A) = \exp(A - A) = \exp(O) = I + O + \frac{1}{2!}O + \cdots = I.$$

■

Remark 4.5. In the previous proposition. it is crucially important that A commutes with B, i.e. that $AB - BA = 0$, for the result to hold true. We would like to understand what happens when A and B do not necessarily commute. This will lead us to the Baker-Campbell-Hausdorff theorem and will provide one motivation for defining the Lie Bracket. Before this we have to define the logarithm function.

We are going to show that the exponential function is locally invertible near the origin and this will enable us to define the logarithm function. We first recall the Inverse Function Theorem from Multivariable Calculus.

Theorem 4.6 (Inverse Function Theorem)**.** *Suppose $F : \mathbb{R}^N \to \mathbb{R}^N$, given in terms of its components as*

$$F(x_1, \dots, x_N) = (F_1(x_1, \dots, x_N), F_2(x_1, \dots, x_n), \dots, F_N(x_1, \dots, x_N)),$$

is a differentiable map such that its derivative D_0F at the origin, given by the Jacobian matrix as

$$D_0F = \begin{pmatrix} \frac{\partial F}{\partial x_1}(0) & \dots & \frac{\partial F_1}{\partial x_N}(0) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1}(0) & \dots & \frac{\partial F_N}{\partial x_N}(0) \end{pmatrix},$$

is invertible (i.e. has nonzero determinant). Then if $y = F(0)$ there exists neighborhoods $0 \in U \subseteq \mathbb{R}^n$ and $y \in V \subseteq \mathbb{R}^N$ such that

$$F_{|V:U\to V}$$

is bijective and F^{-1} is differentiable. Furthermore,

$$D_y(F^{-1}) = (D_O F)^{-1}.$$

(In other words, F is locally invertible near the origin.)

We can now prove the following theorem:

Theorem 4.7. *There exist neighborhoods $O \in U \subseteq M_n(\mathbb{K})$ and $I \in V \subseteq \mathrm{GL}_n(\mathbb{K})$ such that $\exp_{|U} : U \to V$ is bijective (in particular, it has an inverse. The inverse is called* log*).*

Proof. We can think of the exponential map $M_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K})$ as a function from $\mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$. We let $A \in M_n(\mathbb{K})$ and denote

$$A = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix}.$$

and then, viewing the exponential function as a function of the n^2 independent variables $A_{11}, \dots, A_{nn}$, we write

$$\exp(A) = \begin{pmatrix} (\exp(A))_{11} & \dots & (\exp(A))_{1n} \\ \vdots & \ddots & \vdots \\ (\exp(A))_{n1} & \dots & (\exp(A))_{nn} \end{pmatrix}.$$

We then have the Jacobian matrix

$$D_O(\exp) = \begin{pmatrix} \frac{\partial(\exp(A))_{11}}{\partial A_{11}}(0) & \dots & \frac{\partial(\exp(A))_{11}}{\partial A_{nn}}(0) \\ \vdots & \ddots & \vdots \\ \frac{\partial(\exp(A))_{nn}}{\partial A_{11}}(0) & \dots & \frac{\partial(\exp(A))_{nn}}{\partial A_{nn}}(0) \end{pmatrix}.$$

Now recall that we have

$$\begin{pmatrix} (\exp(A))_{11} & \dots & (\exp(A))_{1n} \\ \vdots & \ddots & \vdots \\ (\exp(A))_{n1} & \dots & (\exp(A))_{nn} \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} + \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix}^2 + \dots.$$

Since we are going to differentiate each entry and ultimately put $A_1 = \dots = A_{nn}$, we therefore have

$$D_O(\exp) = \begin{pmatrix} \frac{\partial A_{11}}{\partial A_{11}}(0) & \dots & \frac{\partial A_{11}}{\partial A_{nn}}(0) \\ \vdots & \ddots & \vdots \\ \frac{\partial(A_{nn}}{\partial A_{11}}(0) & \dots & \frac{\partial(A_{nn}}{\partial A_{nn}}(0) \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}.$$

This is invertible. We are now able to use the inverse function theorem near the origin to complete the proof. ■

We now show the following:

Proposition 4.8. *For $\|A\| < 1$, we have*

$$\log(I + A) = A - \frac{1}{2}A^2 + \frac{1}{3}A^3 - \frac{1}{4}A^4 + \dots.$$

Proof. That the infinite series converges when $\|A\| < 1$ (and is continuous then) is proved in a similar way to how we proved the analogous results for exp above. By actual formal manipulations, we can show that

$$\exp\left(\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(A - I)^n\right) = \sum_{m=0}^{\infty} \frac{1}{m!}\left(\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(A - I)^n\right)^m = A$$

and similarly that

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\left(\sum_{m=1}^{\infty} \frac{1}{m!}A^m\right) = A.$$

We now have the required result. ■

It turns out that exp and log are infinitely differentiable, but we will just assume this result and omit the proof for brevity. We have the following result which is analogous to the the case with complex numbers:

Proposition 4.9.

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

Proof.

$$\begin{aligned}\frac{d}{dt}(\exp(tA)) &= \frac{d}{dt}\sum_{n\geq 0}\frac{1}{n!}(tA)^n \\ &= \sum_{n\geq 0}\frac{1}{n!}\frac{d}{dt}(t^nA^n) \\ &= \sum_{n\geq 0}\frac{1}{n!}nt^{n-1}A^n \\ &= A\sum_{n=1}^{\infty}\frac{t^{n-1}}{(n-1)!}A^{n-1} \\ &= A\sum_{n=0}^{\infty}\frac{1}{n!}(tA)^n \\ &= A\exp(tA).\end{aligned}$$

■

We recall that if A and B commute then $\exp(A)\exp(B) = \exp(A+B)$. What happens when $AB \neq BA$? We will use the logarithm to explain what happens. For small A and B we have

$$\begin{aligned}\log\exp(A)\exp(B) &= \log\left(\left(\sum_{n=0}^{\infty}\frac{1}{n!}A^n\right)\left(\sum_{m=0}^{\infty}\frac{1}{m!}B^m\right)\right) \\ &= \log\left(1+A+B+AB+\frac{A^2}{2}+\frac{B^2}{2}+\frac{AB^2}{2}+\frac{A^2B}{2}+\dots\right) \\ &= X-\frac{X^2}{2}+\frac{X^3}{2}-\dots(\text{where } X = A+B+AB+\frac{1}{2}A^2+\frac{1}{2}B^2+\dots) \\ &= A+B+AB+\frac{1}{2}A^2+\frac{1}{2}B^2-\frac{1}{2}(A^2+AB+BA+B^2+\dots)+\dots \\ &= A+B+\frac{1}{2}(AB-BA)+\dots\end{aligned}$$

The term $AB-BA$ is denoted by $[A,B]$ and is called the *commutator bracket* of A and B. It measures the extent to which A and B fail to be commutative. The Baker-Campbell Hausdorff formula, which we will not prove here, states that the third, fourth, … correction terms in the expansion $\log((\exp(A)\exp(B))) = A+B+\frac{1}{2}[A,B]\dots$ can all be expressed in terms of A, B and $[A,B]$, and even gives an explicitly formula for these: (which admittedly is not particularly illuminating)

Theorem 4.10 (Baker-Campbell-Hausdorff Formula)**.** *For small A and B we have*

$$\log((\exp(A)\exp(B))) = \sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}\sum_{\substack{r_1+s_1>0\\ \vdots \\ r_n+s_n>0}}\frac{[A^{(r_1)}B^{(s_1)}A^{(r_2)}B^{(s_2)}\dots A^{(r_n)}B^{(s_n)}\dots\dots]}{\sum_{j=1}^{n}(r_j+s_j)\prod_{i=1}^{n}r_i!s_i!}.$$

Here $[M^{(k)}A] = [M,[M^{(k-1)}A]]$ with $[M^{(1)}A] = [M,A]$.

For example, the cubic correction term is

$$\frac{1}{12}[A,[A,B]] - \frac{1}{12}[B,[A,B]].$$

The main point of the Baker-Campbell-Hausdorff formula is that the group multiplication between the invertible matrices $\exp A$ and $\exp B$ in $\mathrm{GL}_n(\mathbb{K})$ is determined entirely by the bracket operation on A and B as matrices in $M_n(\mathbb{K})$.
We give now an illustrative example of the exponential function in action:

Example. Let $A = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}$. Then it is easy to verify that

$$A^n = \begin{cases} \begin{pmatrix} 0 & -\theta^n \\ \theta^n & 0 \end{pmatrix} & n \equiv 1 \mod 4 \\ \begin{pmatrix} -\theta^n & 0 \\ 0 & -\theta^n \end{pmatrix} & n \equiv 2 \mod 4 \\ \begin{pmatrix} 0 & \theta^n \\ -\theta^n & 0 \end{pmatrix} & n \equiv 3 \mod 4 \\ \begin{pmatrix} \theta^n & 0 \\ 0 & \theta^n \end{pmatrix} & n \equiv 0 \mod 4 \end{cases}$$

so that

$$\exp(A) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

a $2D$-rotation matrix by an angle of θ radians clockwise. Notice we started with an antisymmetric A and exponentiated it to obtain a rotation matrix. Later on we will see that any antisymmetric matrix, when exponentiated gives a rotation matrix.

5. Lie Algebras

So far we have defined an exponential map $\exp : M_n(\mathbb{K}) \to \mathrm{GL}(n,\mathbb{K}))$ such that exp is locally invertible and such that $\exp A \exp B$ can be determined entirely by A, B, and $[,]$. Our goal is to construct a replacement $\mathfrak{g}$ for $M_n(\mathbb{K})$, as well as a replacement G for $\mathrm{GL}(n,\mathbb{K})$, and an exponential map $\exp : \mathfrak{g} \to G$ such that exp has similar properties of local invertibility and such that multiplication of exponentials is determined entirely by a bracket operation on elements of $\mathfrak{g}$. To achieve this, we must introduce an axiomatic definition of a *Lie algebra*:

Definition 5.1 (Lie Algebra)**.** A $\mathbb{K}-$Lie algebra $\mathfrak{g}$ over $\mathbb{K}$ is a vector space over $\mathbb{K}$ equipped with a bilinear map $[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ called the *Lie bracket*, such that for all $x, y, z \in \mathfrak{g}$,

$$[x,y] = -[y,x], \qquad \text{(Skew symmetry)}$$
$$[x,[y,z]] + [y,[x,y]] + [z,[x,y]] = 0. \qquad \text{(Jacobi Identity)}$$

Here $\mathbb{K}-$bilinear means that for all $x_1, x_2, x, y_1, y_2, y \in \mathfrak{a}$ and all $r_1, r_2, r, s_1, s_2, s \in \mathbb{K}$,

$$[r_1x_1 + r_2x_2, y] = r_1[x_1, y] + r_2[x_2, y],$$
$$[x, s_1y_1 + s_2y_2] = s_1[x, y_1] + s_2[x, y_2].$$

Example. Taking $\mathbb{K} = \mathbb{R}$ and $\mathfrak{g} = \mathbb{R}^3$, and taking the Lie bracket to be the cross product $[x,y] = x \times y$ gives a Lie Algebra.

Definition 5.2 (Commutator bracket)**.** Given two matrices $A, B \in M_n(\mathbb{K})$, we define the *commutator bracket* as $[A, B] = AB - BA$.

Proposition 5.3. *The commutator bracket is a Lie bracket.*

Proof. We proceed to directly verify that the commutator bracket satisfies all the properties of a Lie bracket. Let us denote arbitrary scalars by $c_1, c_2 \in \mathbb{K}$ and arbitrary matrices by $A, B, C \in M_n(\mathbb{K})$. First we check $\mathbb{K}-$bilinearity. We have

$$\begin{aligned}[c_1A + c_2B, C] &= (c_1A + c_2B)C - C(c_1A + c_2B)\\ &= c_1(AC - CA) + c_2(BC - CB)\\ &= c_1[A, C] + c_2[B, C].\end{aligned}$$

Similarly we verify that $[C, c_1A + c_2B] = c_1[C, A] + c_2[C, B]$. The skew symmetry is clear

$$[A, B] = AB - BA = -(BA - AB) = -[B, A].$$

Finally we verify the Jacobi identity. We have

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

as desired. ■

Corollary 5.4. *The $\mathbb{K}$-vector space $M_n(\mathbb{K})$ with the commutator bracket is a $\mathbb{K}$-Lie algebra.*

We will now define Lie subalgebras.

Definition 5.5 (Lie subalgebras)**.** A *Lie subalgebra* is a vector subspace $\mathfrak{h} \subseteq \mathfrak{g}$ such that for all $A, B \in \mathfrak{h}$ we have $[A, B] \in \mathfrak{h}$.

We now define paths, curves and tangent spaces.

Definition 5.6 (Path)**.** Let X be a topological space, and $x_0, x_1 \in X$. Then a *path* from x_0 to x_1 is a continuous function $\gamma : [0, 1] \to X$ such that $\gamma(0) : [0, 1] \to X$ such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$.

Definition 5.7 (Derivatives)**.** $\alpha'(t)$ is defined as

$$\alpha'(t) = \lim_{s \to t} \frac{\alpha(s) - \alpha(t)}{s - t} \in M_n(\mathbb{K}),$$

provided this limit exists.

Definition 5.8 (Differentiable Curve)**.** A *differentiable curve* in $M_n(\mathbb{K})$ is a function

$$\alpha : (a, b) \to M_n(\mathbb{K})$$

for which the derivative $\alpha'(t)$ exists for all $t \in (a, b)$.

Definition 5.9 (Tangent Spaces)**.** The tangent space to G at $U \in G$ is

$$T_U G = \{\gamma'(0) \in M_n(k) : \gamma \text{ a differentiable curve in } G \text{ with } \gamma(0) = U\}.$$

We are now going to prove the following important theorem:

Theorem 5.10. *Let G be a matrix subgroup of $\mathrm{GL}_n(\mathbb{K})$. Define $g := \{A \in M_n(\mathbb{K}) : \exp(tA) \in G, \forall t \in \mathbb{R}\}$. Then*

1. *g is a vector subspace of $M_n(\mathbb{K})$.*
2. *If $A, B \in g$, then $[A, B] \in g$.*
3. *g is the tangent space of G at I.*
4. *$\exp : g \to G$ is locally invertible.*

In particular, g is a Lie algebra and is called the Lie algebra of G.

In order to prove the first two claims, we will first prove the third. Before we can do these things we need to introduce an important lemma.

Lemma 5.11. *If $\vec{\gamma}(s)$ is a path in the matrix group G such that $\vec{\gamma}(0) = I$ then $\dot{\vec{\gamma}}(0) \in g$ where $g := \{A \in M_n(\mathbb{K}) : \exp(tA) \in G \forall t \in \mathbb{R}\}$.*

Proof. We must prove that $\exp(t\dot{\vec{\gamma}}(0)) \in G$ for all t. We claim that it suffices to show that $\exp(\dot{\vec{\gamma}}(0)) \in G$; for supposing that $\exp(\dot{\vec{\gamma}}(0)) \in G$ whenever γ satisfies the hypotheses of the lemma, then when we put $\gamma(st) = \delta(s)$ for any t, $\delta(s)$ is a path in G such that $\delta(0) = \gamma(0) = I$. In other words, $\delta(s)$ is a path satisfying the conditions required in the lemma. Therefore, by our assumption, $\exp(\dot{\vec{\delta}}(0)) = \exp(t\dot{\vec{\gamma}}(0)) \in G$, and this is true for any t. We therefore proceed to show that $\exp(\dot{\vec{\gamma}}(0)) \in G$. Let $h(s) = \log \gamma(s)$ for small s. Then

$$\begin{aligned}\frac{dh}{ds}(0) &= (D_I(\log))(\dot{\vec{\gamma}}(0)) = (D_0 \exp)^{-1}(\dot{\vec{\gamma}}(0)) \\ &= \mathrm{Id}_0(\dot{\vec{\gamma}}(0)) = (\dot{\vec{\gamma}}(0)).\end{aligned}$$

It will therefore suffice to show that $\exp(\frac{dh}{ds}(0)) \in G$. Now

$$\frac{dh}{ds}(0) = \lim_{\epsilon \to 0} \frac{h(\epsilon) - h(0)}{\epsilon} = \lim_{\epsilon \to 0} \frac{h(\epsilon)}{\epsilon} = \lim_{n \to \infty} nh(1/n).$$

But now $\gamma(s) = \exp h(s) \in G$ for all small s. Therefore, as G is a group, $\gamma(s)^n = \exp(n(h(s))$ for all positive integers n. Let $s = 1/n$ for sufficiently large n. Therefore $\exp(nh(1/n)) \in G$ for all large n. We will now let $n \to 0$. By continuity of exp,

$$\lim_{n \to \infty} \exp(nh(1/n)) = \exp(\lim_{n \to \infty}(nh(1/n)) = \exp(\frac{dh}{ds}(0)) \in G$$

by the closedness of G. ∎

We make the following defintion:

Definition 5.12 (Tangent vector field)**.** If $\gamma(s)$ is a path in $\mathbb{K}^n$, then $\dot{\vec{\gamma}}$ is called the *tangent vector field* along γ.

We are now in a position to prove Theorem 5.10.

Proof of Theorem 5.10. . Lemma 5.11 implies that g contains all the tangent vectors to G at I. We have therefore shown that the tangent space to the group G at the identity is a subset of g. We will now prove that g *is* the tangent space at the identity of G.

Lemma 5.13. *If $A \in g$, then there is a path $\gamma(s) \in G$ such that $\dot{\vec{\gamma}}(0) = A$ and $\vec{\gamma}(0) = I$.*

Proof. Simply let $\gamma(s) = \exp(sA)$. Then $\dot{\vec{\gamma}}(s) = A\exp(sA)$ so that $\dot{\vec{\gamma}}(0) = A$. Also $\vec{\gamma}(0) = I$. ∎

This completes the third claim. We will now use the latter to prove the first two claims. To prove the first claim, we first show that if $A, B \in g$, then $A + B \in g$. For this, define $\vec{\gamma_1}(s) := \exp(sA)\exp(sB)$. Then $\gamma_1(s) \in G$ for all s. We have $\gamma_1(0) = I$ and $\vec{\gamma_1}(s) = A\exp(sA)\exp(sB) + \exp(sA)B\exp(sB)$ so that $\dot{\vec{\gamma_1}}(0) = X + Y$ and therefore $X + Y \in G$ by the third claim which we have already proved. To complete the proof of the first claim, we remain to show that $\lambda A \in g$ for all $A \in g$ and all $\lambda \in \mathbb{K}$. To achieve this, we let $\vec{\gamma_2}(s) := \exp(s\lambda A)$ so that $\vec{\gamma_2}(0) = I$ and so $\dot{\vec{\gamma_2}}(0) = \lambda x \in G$. Moving on to the second claim, let $\vec{\gamma_3}(s) = \exp(As^{\frac{1}{2}})\exp(Bs^{\frac{1}{2}})\exp(-As^{\frac{1}{2}})\exp(-Bs^{\frac{1}{2}})$. Then

$$
\begin{aligned}
\vec{\gamma_3}(s) &= \exp(As^{\frac{1}{2}} + Bs^{\frac{1}{2}} + \frac{1}{2}s[A,B] + O(s^{\frac{3}{2}}))\exp(-As^{\frac{1}{2}} - Bs^{\frac{1}{2}} + \frac{1}{2}s[A,B] + O(s^{\frac{3}{2}})) \\
&= \exp(s[A,B] - s[A+B, A+B] + O(s^{\frac{3}{2}})) \\
&= \exp(s[A,B] + O(s^{\frac{3}{2}})) \\
\therefore \dot{\vec{\gamma_3}}(0) &= \Big([A+B] + O(s^{\frac{1}{2}})\exp(s[A,B] + O(s^{\frac{3}{2}}))\Big)\Big|_{s=0} = [A,B] \in G
\end{aligned}
$$

completing the proof of the second claim, and thereby establishing that g is indeed a Lie subalgebra of $M_n(\mathbb{K})$. We still have to prove the fourth claim but we omit this for brevity and instead refer the reader to [Eva25]. ■

We will now give two very illustrative examples of matrix groups and their Lie algebras.

Example. We shall find the Lie algebra $\mathfrak{o}_n(\mathbb{R})$ of the orthogonal group $O_n(\mathbb{R})$, which is the group of rotations and reflections in $\mathbb{R}^n$. By Theorem 5.10, we have

$$\mathfrak{o}_n(\mathbb{R}) = \{A : \exp(tA) \in O(n) \forall t \in \mathbb{R}\}.$$

We therefore have

$$
\begin{aligned}
A \in O(n) &\iff (\exp(tA))^T \exp(tA) = I \quad \forall t \in \mathbb{R} \\
&\iff \exp(tX^T)\exp(tX) = I
\end{aligned}
$$

Thus, if $X^T = -X$, i.e. if X is an antisymmetric matrix, then the last condition is satisfied and $X \in \mathfrak{o}_n(\mathbb{R})$. But it is easy to prove that the condition is also necessary because by differentiating both sides we find

$$X^T \exp(tX^T)\exp(tX) + \exp(tX^T)X\exp(tx) = O \quad \forall t.$$

Putting $t = 0$ gives

$$X^T = -X.$$

We have therefore shown that the Lie algebra of the group $O(n)$ of orthogonal matrices is the set of all $n \times n$ antisymmetric matrices.

Example. As the next very important example we shall find the Lie algebra $\mathfrak{sl}_2(\mathbb{R})$ of $\mathrm{SL}_2(\mathbb{R})$. We have, again from Theorem 5.10,

$$\mathfrak{sl}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \det\exp\left(t\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = 1 \quad \forall t \in \mathbb{R} \right\}$$

But now

$$\begin{aligned}\det\exp\left(t\begin{pmatrix}a & b\\ c & d\end{pmatrix}\right) &= \det\left[\begin{pmatrix}1+ta & tb\\ ta & 1+td\end{pmatrix}+O(t^2)\right]\\ &= (1+ta)(1+td)-t^2bc+O(t^2\\ &= 1+t(a+d)+O(d^2)\end{aligned}$$

We must therefore have, for all $t \in \mathbb{R}$,

$$1+t(a+d)+O(d^2)=1.$$

By differentiating both sides and subsequently putting $t = 0$, we find that a sufficient condition is that $a+d$. That is $\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \begin{pmatrix}a & b\\ c & -a\end{pmatrix}$ or, in other words, it suffices for the matrix to be trace-free. We are now going to prove that this condition is also necessary. Consider the set $\left\{\begin{pmatrix}a & b\\ c & -a\end{pmatrix} : a,b,c \in \mathbb{K}\right\}$ of all trace-free matrices. It is easy to see that $\left\{\begin{pmatrix}a & b\\ c & -a\end{pmatrix} : a,b,c \in \mathbb{K}\right\}$ is a three-dimensional vector space under scalar multiplication with elements of $\mathbb{K}$. For basis, we can choose the three vectors $P = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, Q = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$, and $R = \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$. Now we want to show that every member of this set of trace-free matrices is an element of $\mathfrak{sl}_2(\mathbb{R})$. Every member of the last set can be written as $aP+bQ+cR$ for some $a,b,c \in \mathbb{R}$. We want to check that for each trace-free matrix A, we have $\det\exp(tA) = 1$ for all $t \in \mathbb{R}$. It is not difficult to verify that $[P,Q] = 2Q, [P,R] = 2R$ and $[Q,R] = P$. Since $\{P,Q,\}$ is a basis, it follows from the Baker-Campbell-Hausdorff theorem that we need only check that $\det\exp(tP) = \det\exp(tQ) = \det\exp(tR) = 1 \quad \forall t \in \mathbb{R}$. This is easily achieved because one can easily verify that

$$\exp(tP) = \begin{pmatrix}e^t & 0\\ 0 & e^{-t}\end{pmatrix}, \exp(tQ) = \begin{pmatrix}1 & t\\ 0 & 1\end{pmatrix}, \exp(tR) = \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}.$$

We are finally able to conclude that $\mathfrak{sl}_2(\mathbb{R}) = \left\{\begin{pmatrix}a & b\\ c & -a\end{pmatrix} : a,b,c \in \mathbb{R}\right\}$, which is the set the of real trace-free 2×2 matrices.

6. Smooth Manifolds and Tangent Spaces

To introduce smooth manifolds we have to define what it means for a map to be smooth:

Definition 6.1. A continuous map $g : V_1 \to V_2$ where each $V_k \subseteq \mathbb{R}^{m_k}$ is open is *smooth* if it is infinitely differentiable. A smooth bijection g is a *diffeomorphism* if its inverse $g^{-1} : V_2 \to V_1$ is also smooth.

Additionally we need a topological notion:

Definition 6.2. A topological space X is *separable* if it has a countable basis , i.e. a basis of the form $\{U_j\}_{j=1}^{\infty} = \{U_1, U_2, U_3, \dots\}$.

Now for the following definitions we will assume that the topological space M is both separable and Hausdorff.

Definition 6.3. If $U \subseteq M$ and $V \subseteq \mathbb{R}^n$ are open subsets, a homeomorphism $f : U \longrightarrow V$ is called an *n-chart* for U.

If $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ is an open covering of M and $\mathcal{F} = \{f_\alpha : U_\alpha \longrightarrow V_\alpha\}$ is a collection of n-charts, then $\mathcal{F}$ is called an *atlas* for M if, whenever $U_\alpha \cap U_\beta \neq \varnothing$,

$$f_\beta \circ f_\alpha^{-1} : f_\alpha(U_\alpha \cap U_\beta) \longrightarrow f_\beta(U_\alpha \cap U_\beta)$$

is a diffeomorphism.

We will sometimes denote the atlas as $(M,\mathcal{U},\mathcal{F})$ and refer to it as a *smooth manifold of dimension n* or *smooth n-manifold*. We now naturally define what it means to be a submanifold.

Definition 6.4. Let $(M,\mathcal{U},\mathcal{F})$ be a manifold of dimension n. A subset $N \subseteq M$ is a *submanifold of dimension k* if for every $p \in N$ there is an open neighbourhood $U \subseteq M$ of p and an n-chart $f\colon U \longrightarrow V$ such that

$$p \in f^{-1}(V \cap \mathbb{R}^k) = N \cap U.$$

For such an N we can form k-charts of the form

$$f_0\colon N \cap U \longrightarrow V \cap \mathbb{R}^k; \quad f_0(x) = f(x).$$

We will denote this manifold by $(N,\mathcal{U}_N,\mathcal{F}_N)$.

In mathematics we not only want to talk about the objects themselves, but also maps between them. Since we already have a notion of smooth maps from $\mathbb{K}^m$ to $\mathbb{K}^n$, it is natural to want to extend this to manifolds.

Remark 6.5. From here on we will sometimes denote the function composition $f \circ g$ as fg.

Definition 6.6. Let $(M,\mathcal{U},\mathcal{F})$ and $(M',\mathcal{U}',\mathcal{F}')$ be atlases on topological spaces M and M'. A *smooth map* $h : (M,\mathcal{U},\mathcal{F}) \to (M',\mathcal{U}',\mathcal{F}')$ is a continuous map $h : M \to M'$ such that for each pair α, α' with $h(U_\alpha) \cap U'_{\alpha'} \neq \varnothing$, the composite

$$f'_{\alpha'} \circ h \circ f_\alpha^{-1} : f_\alpha(h^{-1}U'_{\alpha'}) \to V'_{\alpha'}$$

is smooth.

In order to talk about Tangent Spaces we first need to talk about curves. Let $(M,\mathcal{U},\mathcal{F})$ be a smooth n-manifold. Let $\gamma : (a,b) \to M$ be a continuous curve with $a < 0 < b$.

Definition 6.7. γ is *differentiable* at $t \in (a,b)$ if for every chart $f : U \longrightarrow V$ with $\gamma(t) \in U$, the curve $f \circ \gamma : (a,b) \longrightarrow V$ is differentiable at $t \in (a,b)$, i.e., $(f \circ \gamma)'(t)$ exists. γ is *smooth* at $t \in (a,b)$ if all the derivatives of $f \circ \gamma$ exist at t.

The curve γ is *differentiable* if it is differentiable at all points in (a,b). Similarly, γ is *smooth* if it is smooth at all points in (a,b).

This presents the issue of how we actually check whether or not a curve is differentiable on a manifold. A curve being differentiable means that it has to be differentiable with respect to every chart. However, we have a lemma that shows that differentiability (and smoothness) of a curve is independent of chart selection.

Lemma 6.8. *Let $f_0 : U_0 \longrightarrow V_0$ be a chart with $\gamma(t) \in U_0$ and suppose that*

$$f_0 \circ \gamma : (a,b) \cap \gamma^{-1} f_0^{-1} V_0 \longrightarrow V_0$$

is differentiable (respectively smooth) at t. Then for any chart $f : U \longrightarrow V$ with $\gamma(t) \in U$,

$$f \circ \gamma : (a,b) \cap \gamma^{-1} f^{-1} V \longrightarrow V$$

is differentiable (respectively smooth) at t.

Proof. Assume that $\gamma(t) \in U_0$ and $\gamma(t)$ is differentiable (respectively smooth) at t with respect to the chart f_0. Now suppose that $\gamma(t) \in U_0 \cap U$ and that f is the chart associated with U. Now we can note by Definition 6.3 that the map

$$\psi = f \circ f_0^{-1} : f_0(U_0 \cap U) \to f(U_0 \to U)$$

is a diffeomorphism. Additionally near t we can say that

$$f \circ \gamma = (f \circ f_0^{-1}) \circ (f_0 \circ \gamma).$$

But this implies that

$$f \circ \gamma = \psi \circ (f_0 \circ \gamma).$$

Thus, $f \circ \gamma$ is a composition of differentiable (respectively smooth) maps and must therefore be differentiable (respectively smooth) at t. ■

We are able to explicitly calculate the derivative $(f \circ \gamma)'(t)$ via the chain rule as follows:

$$(f \circ \gamma)'(t) = \mathrm{Jac}_{f f_0^{-1}}(f_0\gamma(t)) \cdot (f_0 \circ \gamma)'(t). \tag{6.1}$$

Now we have the tools to be able to talk about tangent spaces. In multivariable calculus we defined the tangent space at a point p as the set of all $\gamma'(p)$ where γ is a curve that is differentiable at p. Tangent spaces on manifolds are not all that different. However, we must adjust our definition because a smooth manifold is defined using an atlas.

Let $p \in M$. Suppose the curve γ passes through p, $\gamma(0) = p$, and that γ is differentiable at 0. Then for any chart $f_0 : U_0 \to V_0$ with $\gamma(0) \in U_0$, there is a derivative vector $v_0 = (f\gamma)'(0) \in \mathbb{R}^n$. In passing to another chart $f : U \to V$ with $\gamma(0) \in U$ we can use Equation 6.1 to get

$$(f \circ \gamma)'(0) = \mathrm{Jac}_{f f_0^{-1}}(f_0\gamma(0)) \cdot (f_0 \circ \gamma)'(0).$$

When we define the tangent space T_pM at a point p on a manifold M we are essentially looking at all pairs of the form

$$((f\gamma)'(0)), f : U \to V).$$

where $p \in U$. Now suppose we have $(f_0\gamma)'(0) = v_0$ and $(f_1\gamma)'(0) = v_1$. It makes sense that in T_pM these vectors should be considered equivalent. This is because they are derivatives of the same curve just under different charts. Thus we impose the equivalence relation $\sim$ under which

$$((f_1\gamma)'(0)), f : U_1 \to V_1) \sim ((f_2\gamma)'(0)), f : U_2 \to V_2).$$

Since we have that

$$(f_1\gamma)'(0) = \mathrm{Jac}_{f_2 f_1^{-1}}(f_1\gamma(0)) \cdot (f_1\gamma)'(0),$$

we can represent the relation as

$$(v, f_1 : U_1 \to V_1) \sim (\mathrm{Jac}_{f_2 f_1^{-1}}(f_1(p))v, f_2 : U_2 \to V_2)$$

where there is a curve γ in M such that

$$\gamma(0) = p \qquad \text{and} \qquad (f_1\gamma)'(0) = v.$$

We denote the tangent space T_pM as the set of these equivalence classes. Additionally we sometimes denote the equivalence class $(v, f : U \to V)$ as $[v, f : U \to V]$.

Since we are given this abstract definition of a tangent space it is natural to want to explore its structure. But given this general definition are we even able to say anything concrete? We actually can!

Proposition 6.9. *Suppose we are given a smooth n-manifold* $(M,\mathcal{U},\mathcal{F})$*. For* $p \in M$*,* T_pM *is an* $\mathbb{R}$*-vector space of dimension* n*.*

Proof. It is fairly easy to show that T_pM is a vector space. So we will omit part and focus on show that it is dimension n.

For any chart $f : U \to V$ with $p \in U$ we can identify all elements of T_pM with objects of the form $(v, f : U \to V)$. This means that T_pM is isomorphic to some subset of $\mathbb{R}^n$. If we can show that all vectors in $\mathbb{R}^n$ arise in T_pM we will be done. Take a vector $v \in \mathbb{R}^n$. We can define a curve $\overline{\gamma} : (-\varepsilon, \varepsilon) \to V$ such that $\overline{\gamma}(0) = p$ and

$$\overline{\gamma}(t) = f(p) + tv.$$

Clearly we have $\overline{\gamma}'(0) = v$. Now we define an associated curve γ in M:

$$\gamma : (-\varepsilon, \varepsilon) :\to M; \qquad \gamma(t) = f^{-1}\overline{\gamma}(t).$$

The curve γ satisfies $\gamma(0) = p$ and $(f\gamma)'(0) = \overline{\gamma}'(0) = v$. Thus every $v \in \mathbb{R}^n$ can be identified with the equivalence class

$$[v, f : U \to V] \in T_pM.$$

In other words, T_pM is isomorphic to $\mathbb{R}^n$. ■

Now that we have defined tangent spaces we are able talk about the derivative. Let $h : (M,\mathcal{U},\mathcal{F}) \to (M',\mathcal{U}',\mathcal{F}')$ be a smooth map between manifolds of dimension n, n'. For $p \in M$ we take the pair of charts $f_\alpha, f'_{\alpha'}$ such that $p \in U_\alpha$ and $h(p) \in U'_{\alpha'}$. By Definition 6.6 we have that $h_{\alpha',\alpha} = f'_{\alpha'} \circ h \circ f_\alpha^{-1}$ is smooth. Since it is smooth, we know that the function can be locally approximated by the $\mathbb{R}$-linear transformation that is associated with the Jacobian matrix $\mathrm{Jac}_{h_{\alpha',\alpha}}(f_\alpha(p))\mathbf{x}$ where $\mathbf{x} = (f_\alpha\gamma)'(0)$. Thus we can define the derivative with respect to charts $f_\alpha, f'_{\alpha'}$ as

$$\mathrm{d}\, h_{\alpha',\alpha} : \mathbb{R}^n \to \mathbb{R}^{n'}; \qquad \mathrm{d}\, h_{\alpha',\alpha}(\mathbf{x}) = \mathrm{Jac}_{h_{\alpha',\alpha}}(f_\alpha(p))\mathbf{x}. \tag{6.2}$$

We can verify that this passes to equivalence classes and can thus give a chart independent definition:

Definition 6.10. Let $h : (M,\mathcal{U},\mathcal{F}) \to (M',\mathcal{U}',\mathcal{F}')$ be a smooth map between manifolds. For $p \in M$ the derivative $\mathrm{d}\, h_p$ is defined as follows:

$$\mathrm{d}\, h_p : T_pM \to T_{h(p)}M'; \qquad \mathrm{d}\, h_p([\gamma']) = [(h \circ \gamma)']$$

Most of the time when working with derivatives we don't need anything other than Definition 6.10. A case in which we would need more is when computations are required. In that situation we would apply Equation 6.2. Before moving on to Lie groups we will quickly introduce a theorem which will be used for an example later:

Theorem 6.11 (Implicit Function Theorem for manifolds)**.** *Let $h\colon (M,\mathcal{U},\mathcal{F}) \longrightarrow (M',\mathcal{U},\mathcal{F}')$ be a smooth map between manifolds of dimensions n, n'. Suppose that for some $q \in M'$, $dh_p\colon T_pM \longrightarrow T_{h(p)}M'$ is surjective for every $p \in N = h^{-1}q$. Then $N \subseteq M$ is a submanifold of dimension $n - n'$ and the tangent space at $p \in N$ is given by $T_pN = \ker dh_p$.*

Proof. The proof follows from the Implicit Function Theorem of Calculus. ■

7. Lie Groups

Now we are finally able to define Lie Groups.

Definition 7.1. Let G be a smooth manifold which is also a topological group with multiplication map mult : $G \times G \to G$ and inverse map inv : $G \to G$ and view $G \times G$ as the product manifold. Then G is a Lie group if mult, inv are both smooth maps.

Naturally we also define what it means to be a Lie subgroup:

Definition 7.2. Let G be a Lie Group. A closed subgroup $H \subseteq G$ that is also a submanifold is called *Lie subgroup* of G. It is then automatic that the restriction to H of the multiplication and inverse maps on G are smooth, hence H is also a Lie Group.

As it's important to study Lie Groups through examples we will present one here:

Example. When $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, $\mathrm{GL}_n(\mathbb{K})$ is a Lie Group.

Proof. From Proposition 2.11 we have that $\mathrm{GL}_n(\mathbb{K})$ is an open set. Since we know that $M_n(\mathbb{K})$ is a separable topological space (with the usual topology defined by the matrix norm), we must have that $\mathrm{GL}_n(\mathbb{K})$ is a separable space as well. To define a smooth manifold we also need charts. For the charts we take an open set $U \subseteq \mathrm{GL}_n(\mathbb{K})$ and the identity function $\mathrm{Id} : U \to U_{\mathbb{K}^{n^2}}$ where $U_{\mathbb{K}^{n^2}}$ is U identified with $\mathbb{K}^{n^2}$. Finally, the multiplication and inverse maps are obviously smooth since they are defined by polynomial and rational functions. ■

We will now give an example of a Lie subgroup. However, we first need to introduce a lemma which will be useful.

Lemma 7.3. *Given a smooth curve with $\alpha : (-\varepsilon, \varepsilon) \to GL_n(\mathbb{K})$ and $\alpha(0) = I$ we must have that*

$$\frac{d(\det \alpha(t))}{dt}\bigg|_{t=0} = \operatorname{tr} \alpha'(0).$$

Proof. Recall that the trace of a matrix $A \in M_n(\mathbb{K})$ is

$$\operatorname{tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

Now let $\partial = \frac{d}{dt}|_{t=0}$. It is fairly easy to verify that ∂ has the following property:

$$\partial(\gamma_1\gamma_2) = (\partial\gamma_1)\gamma_2(0) + \gamma_1(0)(\partial\gamma_2).$$

Now let $a_{ij} = \alpha(t)_{ij}$. Then we can notice that when $t = 0$ we have that $a_{ij} = \delta_{ij}$ where δ is the Kronecker Delta function. We now let C_{ij} be the cofactor matrix obtained from $\alpha(t)$ by deleting the ith row and jth column. Using a Laplace expansion we get

$$\det \alpha(t) = \sum_{j=1}^{n} (-1)^{n+j} a_{nj} \det C_{nj}.$$

Then applying then taking the partial on both sides we get

$$\begin{aligned}\partial \det \alpha(t) &= \sum_{j=1}^{n} (-1)^{i+j}((\partial a_{nj}) \det C_{nj} + a_{nj}(\partial \det C_{nj})) \\ &= \sum_{j=1}^{n} (-1)^{i+j}(\partial a_{nj}) \det C_{nj} + \partial \det C_{nn}.\end{aligned}$$

For $t = 0$, we can notice that $\det C_{nj} = \delta_{nj}$. This is clear since if you delete anything other than the nth column you will end up with a column of zeros making the matrix non-invertible. Combining this fact with $\alpha(0) = I$ we get

$$\partial \det(\alpha(t)) = \partial a_{nn} + \partial \det C_{nn}.$$

We can simply repeat this calculation with the $(n-1) \times (n-1)$ matrix C_{nn}. This gives us the following:

$$\begin{aligned}\partial \det(\alpha(t)) &= \partial a_{nn} + \partial a_{(n-1)(n-1)} + \partial \det C_{(n-1)(n-1)} \\ &\vdots \\ &= \partial a_{nn} + \partial a_{(n-1)(n-1)} + \cdots + \partial a_{11} \\ &= \text{tr } \alpha'(0).\end{aligned}$$

■

Using Lemma 7.3 in conjunction with Theorem 6.11 we are able to prove the following statement:

Example. For $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, $\mathrm{SL}_n(\mathbb{K}) \subseteq \mathrm{GL}_n(\mathbb{K})$ is a Lie subgroup.

Proof. We will take $\det : \mathrm{GL}_n(\mathbb{K}) \to \mathbb{K}$ as our smooth function. We already know that $\mathrm{SL}_n(\mathbb{K}) = \det^{-1} 1 \subseteq \mathrm{GL}_n(\mathbb{K})$. Thus to use Theorem 6.11 we only need to show that $\mathrm{d}\ \det_A : M_n(\mathbb{K}) \to \mathbb{K}$ is surjective for all $A \in \mathrm{GL}_n(\mathbb{K})$ (hence making it surjective for all $A \in \mathrm{SL}_n(\mathbb{K})$). To do this we will consider the smooth curve $\alpha : (-\varepsilon, \varepsilon) \to \mathrm{GL}_n(\mathbb{K})$ with $\alpha(0) = A$. The derivative $\mathrm{d}\ \det_A$ applied to $\alpha'(0)$ can be found by

$$d \det_A(\alpha'(0)) = \left.\frac{d \det \alpha(t)}{dt}\right|_{t=0}.$$

We can notice that $\alpha_0(0) = I$. Thus we can apply Lemma 7.3 to get

$$d \det_I(\alpha_0'(0)) = \left.\frac{d \det \alpha_0(t)}{dt}\right|_{t=0} = \text{tr } \alpha_0'(0). \tag{7.1}$$

We now take a modified curve α_0 which is defined as follows:

$$\alpha_0 : (-\varepsilon, \varepsilon) \to \mathrm{GL}_n(\mathbb{K}); \qquad \alpha_0(t) = A^{-1}\alpha(t).$$

Therefore using Equation 7.1 and that $\alpha(t) = A\alpha_0(t)$, we get

$$d \det_A(\alpha'(0)) = \left.\frac{d \det(A\alpha_0(t))}{dt}\right|_{t=0} = \det A \left.\frac{d \det(\alpha_0(t))}{dt}\right|_{t=0} = \det A \text{ tr } \alpha_0'(0).$$

Thus $d \det_A$ is really the $\mathbb{K}$-linear transformation

$$d \det_A : M_n(\mathbb{K}) \to \mathbb{K}; \qquad d \det_A(X) = (\det A)\mathrm{tr}(A^{-1}X).$$

Since the trace function is surjective we must have that $d\det_A$ is surjective as well. Since this is true for all $A \in \mathrm{SL}_n(\mathbb{K})$ we can apply Theorem 6.11 to get that $\mathrm{SL}_n(\mathbb{K})$ is a submanifold of $\mathrm{GL}_n(\mathbb{K})$, thus making $\mathrm{SL}_n(\mathbb{K})$ a Lie subgroup. ∎

In the above example we saw that the matrix subgroup $\mathrm{SL}_n(\mathbb{K}) \subseteq \mathrm{GL}_n(\mathbb{K})$ is also a Lie subgroup of $\mathrm{GL}_n(\mathbb{K})$. It is natural to ask whether this is true in general. Given a matrix subgroup of $\mathrm{GL}_n(\mathbb{K})$ is it true that it is also a Lie subgroup? Amazingly it is!

Theorem 7.4. *Let $G \subseteq \mathrm{GL}_n(\mathbb{K})$ be a matrix subgroup. Then G is a Lie subgroup of $\mathrm{GL}_n(\mathbb{K})$.*

For the sake of brevity we won't prove this, but the proof can be found in Chapter 7 of [Bak02]. Since all Lie subgroups are Lie groups themselves, Theorem 7.4 implies that all matrix groups are Lie groups!

Noting that matrix groups are Lie groups, we can notice a very elegant connection between Lie groups and Lie algebras:

Theorem 7.5. *Let M be a matrix subgroup of $GL_n(\mathbb{K})$. For all $p \in M$, the tangent space T_pM is isomorphic to the Lie algebra on T_eM where e is the identity element of M.*

Proof. Since we've already done most of the work for this proof it will be easy. We first note that Theorem 5.10 states that the tangent space T_eM is in fact a Lie algebra. From there we use Proposition 6.9 to see that all the tangent spaces are of the same dimension and must thus be isomorphic. ∎

This theorem gives us a very interesting connection between Lie groups and Lie algebras. However, one might question whether such a statement is true in general. In fact, the general statement is also true!

Theorem 7.6. *All tangent spaces of a Lie group M are isomorphic to the Lie algebra at the identity.*

Proof. The proof is essentially the same as the proof of Theorem 5.10 except we instead use the general exponentiation function. ∎

8. The Rogers-Ramanujan Identities

We would like to end this paper by briefly mentioning the Rogers-Ramanujan Identities in order to illustrate the power of Lie theoretical methods. The Rogers-Ramanujan Identities are one of the most amazing pairs of identities in the whole of mathematics. They are intriguing in the simplicity and sheer elegance of their statement, their connections to deep areas of mathematics [Sil18] and surprisingly, their importance in statistical mechanics [Bax07] and conformal field theory [BMS98].

Theorem 8.1 (Rogers-Ramanujan Identities)**.**

$$1 + \sum_{k=1}^{\infty} \frac{q^{k^2}}{(1-q)\dots(1-q^k)} = \prod_{n=1}^{\infty}(1-q^{5n+1})^{-1}(1-q^{5n+4})^{-1}$$

$$1 + \sum_{k=1}^{\infty} \frac{q^{k(k+1)}}{(1-q)\dots(1-q^k)} = \prod_{n=1}^{\infty}(1-q^{5n+2})^{-1}(1-q^{5n+3})^{-1}$$

The infinite series are known as the Rogers-Ramanujan functions:

Definition 8.2 (Rogers-Ramanujan functions).

$$G(q) := 1 + \sum_{k=1}^{\infty} \frac{q^{k^2}}{(1-q)\dots(1-q^k)}$$
$$H(q) := 1 + \sum_{k=1}^{\infty} \frac{q^{k(k+1)}}{(1-q)\dots(1-q^k)}$$

This pair of identities was found first by Leonard James Rogers and later independently rediscovered by Srinivasa Ramanujan. A proof was published jointly by Rogers and Ramanujan in [RR19]. However, in order to understand the connection to conformal field theory, one should likely consider the much more difficult Lie theoretical proof [LW82]. A summary of the method of proof can be found at [LW81]. Andrews [And89] also explains the main parts of the proof. Lepowsky and Wilson [LW82] build on work by Lepowsky (a Lie theorist) and Milne (a combinatorialist) [LM78a, LM78b] in order to achieve the goal. Lepowsky and Milne considered the infinite product parts of the Rogers-Ramanujan identities while Lepowsky and Milne considered the infinite series portions. The tools involved include Euclidean Kac-Moody Lie algebras and vertex operators. In the end, the summands in the Rogers-Ramanujan identities turn out “count” dimensions of certain spaces. One can read more about the representation theory of affine Lie algebras in connection to the Rogers-Ramanujan identities in [Isa25]. For an introduction to infinite dimensional Lie algebras, one can read [Kac83]. It should be noted that the theory of vertex operator algebras developed by Lepowsky and Wilson turned out to be highly influential, as it was used to construct a natural representation of the Monster finite simple group in 1988 and was key in the work of Borcherds on vertex algebras and his resolution of the Conway-Norton monstrous moonshine conjecture in 1992.

Acknowledgements

The authors would like to thank Simon-Rubinstein Salzedo for providing the opportunity to write this paper and Jacob Swenberg for reading an earlier version of it.

References

[And89] George E. Andrews. On the proofs of the Rogers-Ramanujan identities. q-Series and partitions, Proc. Workshop, Minneapolis/MN (USA) 1988, IMA Vol. Math. Appl. 18, 1-14 (1989)., 1989.

[Bak02] Andrew Baker. *Matrix groups. An introduction to Lie group theory.* Springer Undergrad. Math. Ser. London: Springer, 2002.

[Bax07] R.J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Dover books on physics. Dover Publications, 2007.

[BMS98] Alexander Berkovich, Barry M McCoy, and Anne Schilling. Rogers–Schur–Ramanujan type identities for the minimal models of conformal field theory. *Communications in mathematical physics*, 191(2):325–395, 1998.

[Bor25] Richard Ewen Borcherds. Lie groups. `https://www.youtube.com/watch?v=pAuRWd8dpvE&list=PL8yHsr3EFj53RWBkiHKoOsTw-dGHAoJ-h`, August 2025.

[Con25] Keith Conrad. Equivalence of Norms. `https://kconrad.math.uconn.edu/blurbs/gradnumthy/equivnorms.pdf`, August 2025.

[Eva25] Jonathan David Evans. Lie groups and Lie algebras. `https://jde27.uk/lgla/`, August 2025.

[Isa25] Konan Isaac. Rogers-Ramanujan type identities: bijectives proofs and Lie- theoretic approach. `https://www.irif.fr/_media/users/konan/konan_isaac_va2.pdf`, August 2025.

[Kac83] Victor G. Kac. *Infinite dimensional Lie algebras. An introduction*, volume 44 of *Prog. Math.* Birkhäuser, Cham, 1983.

[LM78a] J. Lepowsky and S. Milne. Lie algebraic approaches to classical partition identities. *Adv. Math.*, 29:15–59, 1978.

[LM78b] J. Lepowsky and S. Milne. Lie algebras and classical partition identities. *Proc. Natl. Acad. Sci. USA*, 75:578–579, 1978.

[LW81] James Lepowsky and Robert Lee Wilson. The Rogers-Ramanujan identities: Lie theoretic interpretation and proof. *Proc. Natl. Acad. Sci. USA*, 78:699–701, 1981.

[LW82] James Lepowsky and Robert Lee Wilson. A Lie theoretic interpretation and proof of the Rogers-Ramanujan identities. *Adv. Math.*, 45:21–72, 1982.

[RR19] L. J. Rogers and S. Ramanujan. Proof of certain identities in combinatory analysis. *Proc. Camb. Philos. Soc.*, 19:211–216, 1919.

[Sil18] Andrew V. Sills. *An invitation to the Rogers-Ramanujan identities.* Boca Raton, FL: CRC Press, 2018.