## Chapter 4

## Matrix Groups

## Topics :

1. Real and Complex Matrix Groups
2. Examples of Matrix Groups
3. The Exponential Mapping
4. Lie Algebras for Matrix Groups
5. More Properties of the Exponential Mapping
6. Examples of Lie Algebras of Matrix Groups

Copyright © Claudiu C. Remsing, 2006.
All rights reserved.

### 4.1 Real and Complex Matrix Groups

Throughout, we shall denote by $\mathbb{k}$ either the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers.

## The algebra of $n \times n$ matrices over $\mathbb{k}$

Let $\mathbb{k}^{m}$ be the set of all $m$-tuples of elements of $\mathbb{k}$. Under the usual addition and scalar multiplication, $\mathbb{k}^{m}$ is a vector space over $\mathbb{k}$. The set $\operatorname{Hom}\left(\mathbb{k}^{n}, \mathbb{k}^{m}\right.$ ) of all linear mappings from $\mathbb{k}^{n}$ to $\mathbb{k}^{m}$ (i.e., mappings $L: \mathbb{k}^{n} \rightarrow$ $\mathbb{k}^{m}$ such that $L(\lambda x+\mu y)=\lambda L(x)+\mu L(y)$ for every $x, y \in \mathbb{k}^{n}$ and $\left.\lambda, \mu \in \mathbb{k}\right)$ is also a vector space over $\mathbb{k}$.
$\diamond$ Exercise 171 Determine the dimension of the vector space $\operatorname{Hom}\left(\mathbb{k}^{n}, \mathbb{k}^{m}\right)$.
Let $\mathbb{k}^{m \times n}$ be the set of all $m \times n$ matrices with elements (entries) from $\mathbb{k}$. It is convenient to identify
the $m$-tuple $\left(a_{1}, a_{2}, \ldots, a_{m}\right) \in \mathbb{k}^{m}$ with the column m-matrix $\left[\begin{array}{c}a_{1} \\ a_{2} \\ \vdots \\ a_{m}\end{array}\right] \in \mathbb{k}^{m \times 1}$.
$\diamond$ Exercise 172 Give reasons why the identification of $\mathbb{k}^{m}$ with $\mathbb{k}^{m \times 1}$ is legitimate.

Under the usual matrix addition and multiplication, $\mathbb{k}^{m \times n}$ is a vector space over $\mathbb{k}$. There is a natural one-to-one correspondence

$$
A \mapsto L_{A}(: x \mapsto A x)
$$

between the $m \times n$ matrices with elements from $\mathbb{k}$ and the linear mappings from $\mathbb{k}^{n}$ to $\mathbb{k}^{m}$.
$\diamond$ Exercise 173 Show that the vector spaces $\mathbb{k}^{m \times n}$ and Hom $\left(\mathbb{k}^{n}, \mathbb{k}^{m}\right)$ are isomorphic. Observe that, in particular, the vector spaces $\mathbb{k}^{1 \times n}$ and $\operatorname{Hom}\left(\mathbb{k}^{n}, \mathbb{k}\right)=$ $\left(\mathbb{k}^{n}\right)^{*}$ (the dual of $\mathbb{k}^{n}$ ) are isomorphic.

Note : We do not identify the row $n$-matrix $\left[\begin{array}{llll}a_{1} & a_{2} & \ldots & a_{n}\end{array}\right]$ with the $n$-tuple $\left(a_{1}, a_{2}, \ldots, a_{n}\right)$ but rather with the linear mapping (functional)

$$
\left(x_{1}, x_{2}, \ldots, x_{n}\right) \mapsto a_{1} x_{1}+a_{2} x_{2}+\cdots+a_{n} x_{n}
$$

Any matrix $A \in \mathbb{k}^{m \times n}$ can be interpreted as a linear mapping $L_{A} \in$ Hom $\left(\mathbb{k}^{n}, \mathbb{k}^{m}\right)$, whereas any linear mapping $L \in \operatorname{Hom}\left(\mathbb{k}^{n}, \mathbb{k}^{m}\right)$ can be realized as a matrix $A \in \mathbb{k}^{m \times n}$. Henceforth we shall not distinguish notationwise between a matrix $A$ and its corresponding linear mapping $x \mapsto A x$.

Note : A matrix (or linear mapping, if one prefers) $A \in \mathbb{k}^{n \times n}$ can be viewed as a vector field (on $\mathbb{k}^{n}$ ) : $A$ associates to each point $p$ in $\mathbb{k}^{n}$ the tangent vector $A(p)=A p \in \mathbb{k}^{n}$. We may think of a fluid in motion, so that the velocity of the fluid particles passing through $p$ is always $A(p)$. The vector field is then the current of the flow and the paths of the fluid particles are the trajectories. This kind of flow is, of course, very special : $A(p)$ is independent of time, and depends linearly on $p$.

Notice that $\mathbb{k}^{n \times n}$ is not just a vector space. It also has a multiplication which is associative and distributes over addition (on either side). In other words, under the usual addition and multiplication, $\mathbb{K}^{n \times n}$ is a ring (in general not commutative), with identity $I_{n}$. Moreover, for all $A, B \in \mathbb{k}^{n \times n}$ and $\lambda \in \mathbb{k}$,

$$
\lambda(A B)=(\lambda A) B=A(\lambda B) .
$$

Such a structure is called an (associative) algebra over $\mathbb{k}$.
$\diamond$ Exercise 174 Give the definition of an algebra over (the field) $\mathfrak{k}$. Write down all the axioms.

## The topology of $\mathbb{k}^{n \times n}$

For $x \in \mathbb{k}^{n}\left(=\mathbb{k}^{n \times 1}\right)$, let

$$
\|x\|_{2}:=\sqrt{\left|x_{1}\right|^{2}+\left|x_{2}\right|^{2}+\cdots+\left|x_{n}\right|^{2}}
$$

be the 2 -norm (or Euclidean norm) on $\mathbb{k}^{n}$.
Note: For $r \geq 1$, the $r$-norm of $x \in \mathbb{k}^{n}$ is defined as

$$
\|x\|_{r}:=\left(\left|x_{1}\right|^{r}+\left|x_{2}\right|^{r}+\cdots+\left|x_{n}\right|^{r}\right)^{1 / r} .
$$

The following properties hold (for $x, y \in \mathbb{k}^{n}$ and $\lambda \in \mathbb{k}$ ):

$$
\begin{aligned}
& \|x\|_{r} \geq 0, \quad \text { and } \quad\|x\|_{r}=0 \Longleftrightarrow x=0 ; \\
& \|\lambda x\|_{r}=|\lambda|\|x\|_{r} ; \\
& \|x+y\|_{r} \leq\|x\|_{r}+\|y\|_{r} .
\end{aligned}
$$

In practice, only three of the $r$-norms are used, and they are :

$$
\begin{aligned}
\|x\|_{1} & =\left|x_{1}\right|+\left|x_{2}\right|+\cdots+\left|x_{n}\right| \quad \text { (the grid norm) } \\
\|x\|_{2} & =\sqrt{\left|x_{1}\right|^{2}+\left|x_{2}\right|^{2}+\cdots+\left|x_{n}\right|^{2}} \quad \text { (the Euclidean norm); } \\
\|x\|_{\infty}=\lim _{r \rightarrow \infty}\|x\|_{r} & =\max \left\{\left|x_{1}\right|,\left|x_{2}\right|, \ldots,\left|x_{n}\right|\right\} \quad \text { (the max norm) }
\end{aligned}
$$

For $x \in \mathbb{k}^{n}$, we have

$$
\|x\|_{\infty} \leq\|x\|_{2} \leq\|x\|_{1} \leq \sqrt{n} \cdot\|x\|_{2} \leq n \cdot\|x\|_{\infty}
$$

and so any two of these norms are equivalent (i.e., the associated metric topologies are identical). In fact, all norms on a finite dimensional vector space (over $\mathbb{k}$ ) are equivalent.

The metric topology induced by (the Euclidean distance)

$$
(x, y) \mapsto\|x-y\|_{2}
$$

is the natural topology on the set (vector space) $\mathbb{k}^{n}$.
$\diamond$ Exercise 175 Show that, for $x, y \in \mathbb{k}^{n}$,

$$
\left|\|x\|_{2}-\|y\|_{2}\right| \leq\|x-y\|_{2} .
$$

Hence deduce that the function

$$
\|\cdot\|_{2}: \mathbb{k}^{n} \rightarrow \mathbb{R}, \quad x \mapsto\|x\|_{2}
$$

is continuous (with respect to the natural topologies on $\mathbb{k}^{n}$ and $\mathbb{R}$ ).
$\diamond$ Exercise 176 Given $A \in \mathbb{k}^{n \times n}$, show that the linear mapping (on $\mathbb{k}^{n}$ ) $x \mapsto$ $A x$ is continuous (with respect to the natural topology on $\mathbb{k}^{n}$ ).

Let $A \in \mathbb{k}^{n \times n}$. The 2 -norm $\|\cdot\|_{2}$ on $\mathbb{k}^{n \times 1}$ induces a (matrix) norm on $\mathbb{k}^{n \times n}$ by setting

$$
\|A\|:=\max _{\|x\|_{2}=1}\|A x\|_{2} .
$$

The subset $K=\left\{x \in \mathbb{k}^{n} \mid\|x\|_{2}=1\right\} \subset \mathbb{k}^{n}$ is closed and bounded, and so is compact. [A subset of the metric space $\mathbb{K}^{n}$ is compact if and only if it is closed and bounded.] On the other hand, the function $f: K \rightarrow \mathbb{R}, \quad x \mapsto\|A x\|_{2}$ is continuous. [The composition of two continuous maps is a continuous map.] Hence the maximum value $\max _{x \in K}\|A x\|_{2}$ must exist.

Note : The following topological result holds : If $K \subset \mathbb{K}^{n}$ is a (nonempty) compact set, then any continuous function $f: K \rightarrow \mathbb{R}$ is bounded; that is, the image set $f(K)=\{f(x) \mid x \in K\} \subseteq \mathbb{R}$ is bounded. Moreover, $f$ has a global maximum (and a global minimum).
$\diamond$ Exercise 177 Show that the induced norm $\|\cdot\|$ is compatible with its underlying norm $\|\cdot\|_{2}$; that is (for $A \in \mathbb{k}^{n \times n}$ and $x \in \mathbb{k}^{n}$ ),

$$
\|A x\|_{2} \leq\|A\|\|x\|_{2}
$$

$\|\cdot\|$ is a matrix norm on $\mathbb{k}^{n \times n}$, called the operator norm; that is, it has the following four properties (for $A, B \in \mathbb{k}^{n \times n}$ and $\lambda \in \mathbb{k}$ ):
(MN1) $\quad\|A\| \geq 0, \quad$ and $\quad\|A\|=0 \Longleftrightarrow A=0$;
(MN2) $\quad\|\lambda A\|=|\lambda|\|A\|$;
(MN3) $\quad\|A+B\| \leq\|A\|+\|B\|$;
(MN4) $\quad\|A B\| \leq\|A\|\|B\|$.

Note : There is a simple procedure (well-known in numerical linear algebra) for calculating the operator norm of an $n \times n$ matrix $A$. This is

$$
\|A\|=\sqrt{\lambda_{\max }}
$$

where $\lambda_{\max }$ is the largest eigenvalue of the matrix $A^{*} A$. Here $A^{*}$ denotes the Hermitian conjugate (i.e., the conjugate transpose) matrix of $A$; in the case $\mathbb{k}=\mathbb{R}$, $A^{*}=A^{\top}$.

We define a metric $\rho$ on (the algebra) $\mathbb{k}^{n \times n}$ by

$$
\rho(A, B):=\|A-B\| .
$$

Associated to this metric is a natural topology on $\mathbb{k}^{n \times n}$. Hence fundamental topological concepts, like open sets, closed sets, compactness, connectedness, as well as continuity, can be introduced. In particular, we can speak of continuous functions from $\mathbb{k}^{n \times n}$ into $\mathbb{k}$.

Exercise 178 For $1 \leq i, j \leq n$, show that the coordinate function

$$
\operatorname{coord}_{i j}: \mathbb{k}^{n \times n} \rightarrow \mathbb{k}, \quad A \mapsto a_{i j}
$$

is continuous. [Hint : Show first that $\left|a_{i j}\right| \leq\|A\|$ and then verify the defining condition for continuity.]

It follows immediately that if $f: \mathbb{k}^{n^{2}} \rightarrow \mathbb{k}$ is continuous, then the associated function

$$
\widetilde{f}=f \circ\left(\operatorname{coord}_{i j}\right): \mathbb{k}^{n \times n} \rightarrow \mathbb{k}, \quad A \mapsto f\left(\left(a_{i j}\right)\right)
$$

is also continuous. Here $\left(a_{i j}\right)=\left(a_{11}, a_{12}, \ldots, a_{1 n}, a_{21}, \ldots, a_{n n}\right) \in \mathbb{k}^{n^{2}}$.
$\diamond$ Exercise 179 Show that the determinant function

$$
\operatorname{det}: \mathbb{k}^{n \times n} \rightarrow \mathbb{k}, \quad A \mapsto \operatorname{det} A:=\sum_{\sigma \in S_{n}}(-1)^{|\sigma|} a_{1 \sigma(1)} a_{2 \sigma(2)} \cdots a_{n \sigma(n)}
$$

and the trace function

$$
\operatorname{tr}: \mathbb{k}^{n \times n} \rightarrow \mathbb{k}, \quad A \mapsto \operatorname{tr} A:=\sum_{i=1}^{n} a_{i i}
$$

are continuous.
The metric space $\left(\mathbb{k}^{n \times n}, \rho\right)$ is complete. This means that every Cauchy sequence $\left(A_{r}\right)_{r \geq 0}$ in $\mathbb{k}^{n \times n}$ has a unique limit $\lim _{r \rightarrow \infty} A_{r}$. Furthermore,

$$
\left(\lim _{r \rightarrow \infty} A_{r}\right)_{i j}=\lim _{r \rightarrow \infty}\left(A_{r}\right)_{i j}
$$

Indeed, the limit on the RHS exists, so it is sufficient to check that the required matrix limit is the matrix $A$ with $a_{i j}=\lim _{r \rightarrow \infty}\left(A_{r}\right)_{i j}$. The sequence $\left(A_{r}-A\right)_{r \geq 0}$ satisfies

$$
\left\|A_{r}-A\right\| \leq \sum_{i, j=1}^{n}\left|\left(A_{r}\right)_{i j}-a_{i j}\right| \rightarrow 0 \quad \text { as } r \rightarrow \infty
$$

and so $A_{r} \rightarrow A$.

## Groups of matrices

Let $\operatorname{GL}(n, \mathbb{k})$ be the set of all invertible $n \times n$ matrices over $\mathbb{k}$ (or, equivalently, the set of all linear transformations on $\mathbb{k}^{n}$ ). So

$$
\mathrm{GL}(n, \mathbb{k}):=\left\{A \in \mathbb{k}^{n \times n} \mid \operatorname{det} A \neq 0\right\} .
$$

$\diamond$ Exercise 180 Verify that the set $\mathrm{GL}(n, \mathbb{k})$ is a group under matrix multiplication.
$G L(n, \mathbb{k})$ is called the general linear group over $\mathbb{k}$. We will refer to $\mathrm{GL}(n, \mathbb{R})$ and $\mathrm{GL}(n, \mathbb{C})$ as the real and complex general linear group, respectively.

A $1 \times 1$ matrix over $\mathbb{k}$ is just an element of $\mathbb{k}$ and matrix multiplication of two such elements is just multiplication in $\mathbb{k}$. So we see that

$$
\mathrm{GL}(1, \mathbb{k})=\mathbb{k}^{\times} \quad(\text { the multiplicative group of } \mathbb{k} \backslash\{0\})
$$

4.1.1 Proposition. $\quad G L(n, \mathbb{k})$ is an open subset of $\mathbb{k}^{n \times n}$.

Proof : We have seen that the function det : $\mathbb{k}^{n \times n} \rightarrow \mathbb{k}$ is continuous (see Exercise 179). Then observe that

$$
\mathrm{GL}(n, \mathbb{k})=\mathbb{k}^{n \times n} \backslash \operatorname{det}^{-1}(0)
$$

Since the set $\{0\}$ is closed (in $\mathbb{k})$, it follows that $\operatorname{det}^{-1}(0)=\operatorname{det}^{-1}(\{0\}) \subset$ $\mathbb{k}^{n \times n}$ is also closed. [The preimage of a closed set under a continuous map is a closed set.] Hence $G L(n, \mathbb{k})$ is open. [The complement of a closed set is an open set.]

Let $G$ be a subgroup of the general linear group $\mathrm{GL}(n, \mathbb{k})$. If $G$ is also a closed subspace of $G L(n, \mathbb{k})$, we say that $G$ is a closed subgroup.
4.1.2 Definition. A closed subgroup of $G L(n, \mathbb{k})$ is called a matrix group over $\mathbb{k}$ (or a matrix subgroup of $\operatorname{GL}(n, \mathbb{k})$ ).

Matrix groups are also known as linear groups or even as matrix Lie groups. This latter terminology emphasizes the remarkable fact that every matrix group is a Lie group.

Note : The condition that the group of matrices $G \subseteq G L(n, \mathbb{k})$ is a closed subset of (the metric space) $\mathrm{GL}(n, \mathbb{k})$ means that the following condition is satisfied : if $\left(A_{r}\right)_{r \geq 0}$ is any sequence of matrices in $G$ and $A_{r} \rightarrow A$, then either $A \in G$ or $A$ is not invertible (i.e. $A \notin \mathrm{GL}(n, \mathbb{k})$ ).

The condition that $G$ be a closed subgroup, as opposed to merely a subgroup, should be regarded as a "technicality" since most of the interesting subgroups of $\mathrm{GL}(n, \mathbb{k})$ have this property. Almost all of the matrix groups we will consider have the stronger property that if $\left(A_{r}\right)_{r \geq 0}$ is any sequence of matrices in $G$ converging to some matrix $A$, then $A \in G$.

We will often use the notation $G \leq G L(n, \mathbb{k})$ to indicate that $G$ is a (matrix) subgroup of $\operatorname{GL}(n, \mathbb{k})$.
4.1.3 Example. The general linear group $G L(n, \mathbb{k})$ is a matrix group (over $\mathfrak{k}$ ).
4.1.4 Example. An example of a group of matrices which is not a matrix group is the set of all $n \times n$ invertible matrices all of whose entries are rational numbers. This is in fact a subgroup of $\mathrm{GL}(n, \mathbb{C})$ but not a closed subgroup; that is, one can (easily) have a sequence of invertible matrices with rational entries converging to an invertible matrix with some irrational entries.
$\diamond$ Exercise 181* Let $a \in \mathbb{R} \backslash \mathbb{Q}$. Show that

$$
G=\left\{\left.\left[\begin{array}{cc}
e^{i t} & 0 \\
0 & e^{i a t}
\end{array}\right] \right\rvert\, t \in \mathbb{R}\right\}
$$

is a subgroup of $\mathrm{GL}(2, \mathbb{C})$, and then find a sequence of matrices in $G$ which converges to $-I_{2} \notin G$. This means that $G$ is not a matrix group. [Hint : By taking $t=$ $(2 n+1) \pi$ for a suitably chosen $n \in \mathbb{Z}$, we can make ta arbitrarily close to an odd integer multiple of $\pi,(2 m+1) \pi$ say. It is sufficient to show that for any positive integer $N$, there exist $n, m \in \mathbb{Z}$ such that $|(2 n+1) a-(2 m+1)|<\frac{1}{N}$.]

Note : The closure of $G$ (in $G L(2, \mathbb{C})$ ) can be thought of as (the direct product) $\mathbb{S}^{1} \times \mathbb{S}^{1}$ and so is a matrix group (see Exercise 195).
4.1.5 Proposition. Let $G$ be a matrix group over $\mathbb{k}$ and $H$ a closed subgroup of $G$. Then $H$ is a matrix group over $\mathbb{k}$.

Proof : Every sequence $\left(A_{r}\right)_{r \geq 0}$ in $H$ with a limit in $G L(n, \mathbb{k})$ actually has its limit in $G$ since each $A_{r} \in H \subseteq G$ and $G$ is closed in $\mathrm{GL}(n, \mathbb{k})$. Since $H$ is closed in $G$, this means that $\left(A_{r}\right)_{r \geq 0}$ has a limit in $H$. So $H$ is closed in $\operatorname{GL}(n, \mathbb{k})$, showing it is a matrix group over $\mathbb{k}$.

Exercise 182 Prove that any intersection of matrix groups (over $\mathbb{k}$ ) is a matrix group.
4.1.6 Example. $\quad$ Denote by $\operatorname{SL}(n, \mathbb{k})$ the set of all $n \times n$ matrices over $\mathbb{k}$, having determinant one. So

$$
\mathrm{SL}(n, \mathbb{k}):=\left\{A \in \mathbb{k}^{n \times n} \mid \operatorname{det} A=1\right\} \subseteq \mathrm{GL}(n, \mathbb{k})
$$

$\diamond$ Exercise 183 Show that $\operatorname{SL}(n, \mathbb{k})$ is a closed subgroup of $\operatorname{GL}(n, \mathbb{k})$ and hence is a matrix group over $\mathbb{k}$.
$\operatorname{SL}(n, \mathbb{k})$ is called the special linear group over $\mathbb{k}$. We will refer to $\operatorname{SL}(n, \mathbb{R})$ and $\operatorname{SL}(n, \mathbb{C})$ as the real and complex special linear groups, respectively.
4.1.7 Definition. A closed subgroup of a matrix group $G$ is called a matrix subgroup of $G$.
4.1.8 Example. We can consider $G L(n, \mathbb{k})$ as a subgroup of $G L(n+1, \mathbb{k})$ by identifying the $n \times n$ matrix $A=\left[a_{i j}\right]$ with

$$
\left[\begin{array}{cc}
1 & 0 \\
0 & A
\end{array}\right]=\left[\begin{array}{cccc}
1 & 0 & \ldots & 0 \\
0 & a_{11} & \ldots & a_{1 n} \\
0 & a_{21} & \ldots & a_{2 n} \\
\vdots & \vdots & & \vdots \\
0 & a_{n 1} & \ldots & a_{n n}
\end{array}\right] .
$$

It is easy to verify that $G L(n, \mathbb{k})$ is closed in $G L(n+1, \mathbb{k})$ and hence $G L(n, \mathbb{k})$ is a matrix subgroup of $\operatorname{GL}(n+1, \mathbb{k})$.

Exercise 184 Show that $\operatorname{SL}(n, \mathbb{k})$ is a matrix subgroup of $\operatorname{SL}(n+1, \mathbb{k})$.

### 4.2 Examples of Matrix Groups

The vector space $\mathbb{k}^{n \times n}$ over $\mathbb{k}$ can be considered to be a real vector space, of dimension $n^{2}$ or $2 n^{2}$, respectively. Explicitly, $\mathbb{R}^{n \times n}$ is (isomorphic to) $\mathbb{R}^{n^{2}}$, and $\mathbb{C}^{n \times n}$ is (isomorphic to) $\mathbb{C}^{n^{2}}=\mathbb{R}^{2 n^{2}}$. Hence we may assume, without any loss of generality, that $\mathbb{k}^{n \times n}$ is some Euclidean space $\mathbb{R}^{m}$.

## The real general linear group $G L(n, \mathbb{R})$

We showed that $G L(n, \mathbb{R})$ is a matrix group and that it is an open subset of the vector space $\mathbb{R}^{n \times n}\left(=\mathbb{R}^{n^{2}}\right)$. Since the set $G L(n, \mathbb{R})$ is not closed, it is not compact. [Any compact set is a closed set.]

The determinant function det $: G L(n, \mathbb{R}) \rightarrow \mathbb{R}$ is continuous (in fact, smooth) and maps $G L(n, \mathbb{R})$ onto the two components of $\mathbb{R}^{\times}$. Thus $G L(n, \mathbb{R})$ is not connected. [The image of a connected set under a continuous map is a connected set.]

Note : A matrix group $G$ is said to be connected if given any two matrices $A, B \in G$, there exists a continuous path $\gamma:[a, b] \rightarrow G$ with $\gamma(a)=A$ and $\gamma(b)=B$. This property is what is called path-connectedness in topology, which is not (in general) the same as connectedness. However, it is a fact (not particularly obvious at the moment) that a matrix group is connected if and only if it is path-connected. So in a slight abuse of terminology we shall continue to refer to the above property as connectedness.

A matrix group $G$ which is not connected can be decomposed (uniquely) as a union of several pieces, called components, such that two elements of the same component can be joined by a continuous path, but two elements of different components cannot. The component of $G$ containing the identity is a closed subgroup of $G$ (and hence a connected matrix group).

Consider the sets

$$
\begin{aligned}
\mathrm{GL}^{+}(n, \mathbb{R}) & :=\{A \in \mathrm{GL}(n, \mathbb{R}) \mid \operatorname{det} A>0\} \\
\mathrm{GL}^{-}(n, \mathbb{R}) & :=\{B \in \mathrm{GL}(n, \mathbb{R}) \mid \operatorname{det} B<0\}
\end{aligned}
$$

These two disjoint subsets of $G L(n, \mathbb{R})$ are open and such that

$$
\mathrm{GL}^{+}(n, \mathbb{R}) \cup \mathrm{GL}^{-}(n, \mathbb{R})=\mathrm{GL}(n, \mathbb{R})
$$

[The preimage of an open set under a continuous map is an open set.]
$\diamond$ Exercise 185 Show that $\mathrm{GL}^{+}(n, \mathbb{R})$ is a matrix subgroup of $\mathrm{GL}(n, \mathbb{R})$ but $\mathrm{GL}^{-}(n, \mathbb{R})$ is not.

The mapping

$$
A \in \mathrm{GL}^{+}(n, \mathbb{R}) \mapsto S A \in \mathrm{GL}^{-}(n, \mathbb{R})
$$

where $S=\operatorname{diag}(1,1, \ldots, 1,-1)$, is a bijection (in fact, a diffeomorphism). The transformation $x \mapsto S x$ may be thought of as a reflection in the hyperplane $\mathbb{R}^{n-1}=\mathbb{R}^{n-1} \times\{0\} \subset \mathbb{R}^{n}$.

Note : The group $\mathrm{GL}^{+}(n, \mathbb{R})$ is connected, which proves that $\mathrm{GL}^{+}(n, \mathbb{R})$ is the connected component of the identity in $\mathrm{GL}(n, \mathbb{R})$ and that $\mathrm{GL}(n, \mathbb{R})$ has two (connected) components.

## The real special linear group $\operatorname{SL}(n, \mathbb{R})$

Recall that

$$
\mathrm{SL}(n, \mathbb{R}):=\{A \in \mathrm{GL}(n, \mathbb{R}) \mid \operatorname{det} A=1\}=\operatorname{det}^{-1}(1)
$$

It follows that $\operatorname{SL}(n, \mathbb{R})$ is a closed subgroup of $G L(n, \mathbb{R})$ and hence is a matrix group. [The preimage of a closed set under a continuous map is a closed set.] We introduce a new matrix norm on $\mathbb{R}^{n \times n}$, called the Frobenius norm, as follows :

$$
\|A\|_{F}:=\sqrt{\operatorname{tr}\left(A^{\top} A\right)}=\sqrt{\sum_{i, j=1}^{n} a_{i j}^{2}}
$$

Note : The Frobenius norm coincides with the Euclidean norm on $\mathbb{R}^{n^{2}}$, and is much easier to compute than the operator norm. However, all matrix norms on $\mathbb{R}^{n \times n}$ are equivalent (i.e., they generate the same metric topology).

We shall use this (matrix) norm to show that $\operatorname{SL}(n, \mathbb{R})$ is not compact. Indeed, all matrices of the form

$$
\left[\begin{array}{cccc}
1 & 0 & \ldots & t \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & & \vdots \\
0 & 0 & \ldots & 1
\end{array}\right]
$$

are elements of $\operatorname{SL}(n, \mathbb{R})$ whose norm equals $\sqrt{n+t^{2}}$ for any $t \in \mathbb{R}$. Thus $\mathrm{SL}(n, \mathbb{R})$ is not a bounded subset of $\mathbb{R}^{n \times n}$ and hence is not compact. [In a metric space, any compact set is bounded.]

Note : The special linear group $\operatorname{SL}(n, \mathbb{R})$ is connected.
More on $\operatorname{SL}(2, \mathbb{R})$. $\qquad$

The orthogonal and special orthogonal groups $\mathrm{O}(n)$ and $\mathrm{SO}(n)$

The set

$$
\mathrm{O}(n):=\left\{A \in \mathbb{R}^{n \times n} \mid A^{\top} A=I_{n}\right\}
$$

is the orthogonal group. Clearly, every orthogonal matrix $A \in \mathrm{O}(n)$ has an inverse, namely $A^{\top}$. Hence $\mathrm{O}(n) \subseteq \mathrm{GL}(n, \mathbb{R})$.
$\diamond$ Exercise 186 Verify that $\mathrm{O}(n)$ is a subgroup of the general linear group $\mathrm{GL}(n, \mathbb{R})$.

The single matrix equation $A^{\top} A=I_{n}$ is equivalent to $n^{2}$ equations for the $n^{2}$ real numbers $a_{i j}, i, j=1,2, \ldots, n$ :

$$
\sum_{k=1}^{n} a_{k i} a_{k j}=\delta_{i j} .
$$

This means that $\mathrm{O}(n)$ is a closed subset of $\mathbb{R}^{n \times n}$ and hence of $\mathrm{GL}(n, \mathbb{R})$.
$\diamond$ Exercise 187 Prove that $O(n)$ is a closed subset of $\mathbb{R}^{n^{2}}$.

Thus $\mathrm{O}(n)$ is a matrix group. The group $\mathrm{O}(n)$ is also bounded in $\mathbb{R}^{n \times n}$. Indeed, the (Frobenius) norm of $A \in \mathrm{O}(n)$ is

$$
\|A\|_{F}=\sqrt{\operatorname{tr}\left(A^{\top} A\right)}=\sqrt{\operatorname{tr} I_{n}}=\sqrt{n}
$$

Hence the group $\mathrm{O}(n)$ is compact. [A subset of $\mathbb{R}^{n \times n}$ is compact if and only if it is closed and bounded.]

Let us consider the determinant function (restricted to $\mathrm{O}(n)$ ), det : $\mathrm{O}(n) \rightarrow$ $\mathbb{R}^{\times}$. Then for $A \in \mathrm{O}(n)$

$$
\operatorname{det} I_{n}=\operatorname{det}\left(A^{\top} A\right)=\operatorname{det} A^{\top} \cdot \operatorname{det} A=(\operatorname{det} A)^{2} .
$$

Hence $\operatorname{det} A= \pm 1$. So we have

$$
\mathrm{O}(n)=\mathrm{O}^{+}(n) \cup \mathrm{O}^{-}(n)
$$

where
$\mathrm{O}^{+}(n):=\{A \in \mathrm{O}(n) \mid \operatorname{det} A=1\} \quad$ and $\quad \mathrm{O}^{-}(n):=\{A \in \mathrm{O}(n) \mid \operatorname{det} A=-1\}$.
Note : The group $\mathrm{O}^{+}(n)$ is connected, which proves that $\mathrm{O}^{+}(n)$ is the connected component of the identity in $\mathrm{O}(n)$.

The special orthogonal group is defined as

$$
\mathrm{SO}(n):=\mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R}) .
$$

That is,

$$
\mathrm{SO}(n)=\{A \in \mathrm{O}(n) \mid \operatorname{det} A=1\}=\mathrm{O}^{+}(n) .
$$

It follows that $\mathrm{SO}(n)$ is a closed subset of $\mathrm{O}(n)$ and hence is compact. [A closed subset of a compact set is compact.]

Note : One of the main reasons for the study of these groups $\mathrm{O}(n), \mathrm{SO}(n)$ is their relationship with isometries (i.e., distance-preserving transformations on the Euclidean space $\mathbb{R}^{n}$ ). If such an isometry fixes the origin, then it is actually a linear transformation and so - with respect to the standard basis - corresponds to a matrix $A$. The isometry condition is equivalent to the fact that (for all $x, y \in \mathbb{R}^{n}$ )

$$
A x \bullet A y=x \bullet y,
$$

which in turn is equivalent to the condition that $A^{\top} A=I_{n}$ (i.e., $A$ is orthogonal). Elements of SO ( $n$ ) are (identified with) rotations (or direct isometries); elements of $\mathrm{O}^{-}(n)$ are sometimes referred to as indirect isometries.

## The Lorentz group Lor $(1, n)$

Consider the inner-product (i.e., nondegenerate symmetric bilinear form) $\odot$ on (the vector space) $\mathbb{R}^{n+1}$ given by (for $x, y \in \mathbb{R}^{n+1}$ )

$$
x \odot y:=-x_{1} y_{1}+\sum_{i=2}^{n+1} x_{i} y_{i}
$$

(the so-called Minkowski product). It is standard to denote this inner-product space by $\mathbb{R}^{1, n}$.
$\diamond$ Exercise 188 Show that the group of all linear isometries (i.e., linear transformations on $\mathbb{R}^{1, n}$ that preserve the Minkowski product) is isomorphic to the matrix group

$$
\mathrm{O}(1, n):=\left\{A \in \mathrm{GL}(n+1, \mathbb{R}) \mid A^{\top} S A=S\right\}
$$

where

$$
S=\operatorname{diag}(-1,1,1, \ldots, 1)=\left[\begin{array}{cc}
-1 & 0 \\
0 & I_{n}
\end{array}\right] \in \mathrm{GL}(n+1, \mathbb{R})
$$

In a similar fashion, one can define more general matrix groups

$$
\mathrm{O}(k, \ell) \leq \mathrm{GL}(k+\ell, \mathbb{R}) \quad \text { and } \quad \mathrm{SO}(k, \ell) \leq \mathrm{SL}(k+\ell, \mathbb{R})
$$

usually called "pseudo-orthogonal" groups.

Exercise 189 Define the inner-product $\langle\cdot, \cdot\rangle_{k, \ell}$ on $\mathbb{R}^{k+\ell}$ by the formula

$$
\langle x, y\rangle_{k, \ell}:=-x_{1} y_{1}-\cdots-x_{k} y_{k}+x_{k+1} y_{k+1}+\cdots+x_{k+\ell} y_{k+\ell}
$$

The pseudo-orthogonal group $\mathrm{O}(k, \ell)$ consists of all matrices $A \in \mathrm{GL}(k+\ell, \mathbb{R})$ which preserve this inner-product (i.e., such that $\langle A x, A y\rangle_{k, \ell}=\langle x, y\rangle_{k, \ell}$ for all $x, y \in \mathbb{R}^{k+\ell}$ ).
(a) Verify that $\mathrm{O}(k, \ell)$ is a matrix subgroup of $\mathrm{GL}(k+\ell, \mathbb{R})$.
(b) Let

$$
Q=\operatorname{diag}(-1, \ldots,-1,1, \ldots, 1)=\left[\begin{array}{cc}
-I_{k} & 0 \\
0 & I_{\ell}
\end{array}\right]
$$

Prove that a matrix $A \in \mathrm{GL}(k+\ell, \mathbb{R})$ is in $\mathrm{O}(k, \ell)$ if and only if $A^{\top} Q A=$ $Q$. Hence deduce that $\operatorname{det} A= \pm 1$.
(c) Verify that $\mathrm{SO}(k, \ell):=\mathrm{O}(k, \ell) \cap \mathrm{SL}(k+\ell, \mathbb{R})$ is a matrix subgroup of $\mathrm{SL}(k+\ell, \mathbb{R})$.

Note : Since $O(k, \ell)$ and $O(\ell, k)$ are essentially the same group, we may assume (without any loss of generality) that $1 \leq k \leq \ell$. The pseudo-orthogonal groups are neither compact nor connected. The groups $\mathrm{O}(k, \ell)$ have four (connected) components, whereas $\mathrm{SO}(k, \ell)$ have two components.

For each positive number $\rho>0$, the hyperboloid

$$
\mathcal{H}_{1, n}(\rho):=\left\{x \in \mathbb{R}^{1, n} \mid\langle x, x\rangle=-\rho\right\}
$$

has two (connected) components

$$
\mathcal{H}_{1, n}^{+}(\rho)=\left\{x \in \mathcal{H}_{1, n}(\rho) \mid x_{1}>0\right\} \quad \text { and } \quad \mathcal{H}_{1, n}^{-}(\rho)=\left\{x \in \mathcal{H}_{1, n}(\rho) \mid x_{1}<0\right\}
$$

We define the Lorentz group Lor $(1, n)$ to be the (closed) subgroup of SO $(1, n)$ preserving each of the connected sets $\mathcal{H}_{1, n}^{ \pm}(1)$. Thus

$$
\operatorname{Lor}(1, n):=\left\{A \in \mathrm{SO}(1, n) \mid A \mathcal{H}_{1, n}^{ \pm}(1)=\mathcal{H}_{1, n}^{ \pm}(1)\right\} \leq \mathrm{SO}(1, n)
$$

It turns out that $A \in \operatorname{Lor}(1, n)$ if and only if it preserves the hyperboloids $\mathcal{H}_{1, n}^{ \pm}(\rho), \rho>0$ and the "light cones" $\mathcal{H}_{1, n}^{ \pm}(0)$.

Note : The Lorentz group Lor $(1, n)$ is connected.
Of particular interest in physics is the Lorentz group Lor $=\operatorname{Lor}(1,3)$. That is,

$$
\text { Lor }=\left\{L \in \mathrm{SO}(1,3) \mid L \mathcal{H}_{1,3}^{ \pm}(\rho)=\mathcal{H}_{1,3}^{ \pm}(\rho), \rho \geq 0\right\} \leq \mathrm{SO}(1,3)
$$

$\diamond$ Exercise 190 Show that
(a) The matrix $A=\left[\begin{array}{ll}\cosh t & \sinh t \\ \sinh t & \cosh t\end{array}\right]$ is in $\mathrm{SO}(1,1)$.
(b) For every $s, t \in \mathbb{R}$

$$
\left[\begin{array}{ll}
\cosh s & \sinh s \\
\sinh s & \cosh s
\end{array}\right]\left[\begin{array}{ll}
\cosh t & \sinh t \\
\sinh t & \cosh t
\end{array}\right]=\left[\begin{array}{ll}
\cosh (s+t) & \sinh (s+t) \\
\sinh (s+t) & \cosh (s+t)
\end{array}\right] .
$$

(c) Every element (matrix) of $\mathrm{O}(1,1)$ can be written in one of the four forms

$$
\left[\begin{array}{cc}
\cosh t & \sinh t \\
\sinh t & \cosh t
\end{array}\right],\left[\begin{array}{cc}
-\cosh t & \sinh t \\
\sinh t & -\cosh t
\end{array}\right],\left[\begin{array}{cc}
\cosh t & -\sinh t \\
\sinh t & -\cosh t
\end{array}\right],\left[\begin{array}{cc}
-\cosh t & -\sinh t \\
\sinh t & \cosh t
\end{array}\right] .
$$

(Since $\cosh t$ is always positive, there is no overlap among the four cases. Matrices of the first two forms have determinant one; matrices of the last two forms have determinant minus one.)

Note : We can write

$$
\begin{aligned}
\operatorname{SO}(1,1) & =\operatorname{Lor}(1,1) \cup\left[\begin{array}{cc}
-1 & 0 \\
0 & -1
\end{array}\right] \operatorname{Lor}(1,1) \\
\mathrm{O}(1,1) & =\operatorname{SO}(1,1) \cup\left[\begin{array}{cc}
1 & 0 \\
0 & -1
\end{array}\right] \operatorname{SO}(1,1)
\end{aligned}
$$



## The real symplectic group $\operatorname{Sp}(2 n, \mathbb{R})$

Let

$$
\mathbb{J}:=\left[\begin{array}{cc}
0 & I_{n} \\
-I_{n} & 0
\end{array}\right] \in \operatorname{SL}(2 n, \mathbb{R})
$$

A matrix $A \in \mathbb{R}^{2 n \times 2 n}$ is called symplectic if

$$
A^{\top} \mathbb{J} A=\mathbb{J} .
$$

Note : The word symplectic was invented by Hermann Weyl (1885-1955), who substituted Greek for Latin roots in the word complex to obtain a term which would describe a group (related to "line complexes" but which would not be confused with complex numbers).

Let $\operatorname{Sp}(2 n, \mathbb{R})$ be the set of all $2 n \times 2 n$ symplectic matrices. Taking determinants of the condition $A^{\top} \mathbb{J} A=\mathbb{J}$ gives

$$
1=\operatorname{det} \mathbb{J}=\left(\operatorname{det} A^{\top}\right) \cdot(\operatorname{det} \mathbb{J}) \cdot(\operatorname{det} A)=(\operatorname{det} A)^{2} .
$$

Hence $\operatorname{det} A= \pm 1$, and so $A \in \mathrm{GL}(2 n, \mathbb{R})$. Furthermore, if $A, B \in \operatorname{Sp}(2 n, \mathbb{R})$, then

$$
(A B)^{\top} \mathbb{J}(A B)=B^{\top} A^{\top} \mathbb{J} A B=\mathbb{J} .
$$

Hence $A B \in \operatorname{Sp}(2 n, \mathbb{R})$. Now, if $A^{\top} J A=\mathbb{J}$, then

$$
\mathbb{J} A=\left(A^{\top}\right)^{-1} \mathbb{J}=\left(A^{-1}\right)^{\top} \mathbb{J}
$$

so

$$
\mathbb{J}=\left(A^{-1}\right)^{\top} \mathbb{J} A^{-1} .
$$

It follows that $A^{-1} \in \operatorname{Sp}(2 n, \mathbb{R})$ and hence $\operatorname{Sp}(2 n, \mathbb{R})$ is a group. In fact, it is a closed subgroup of $\mathrm{GL}(2 n, \mathbb{R})$, and thus a matrix group.

Note : The symplectic group $\mathrm{Sp}(2 n, \mathbb{R})$ is connected. (It turns out that the determinant of a symplectic matrix must be positive; this fact is by no means obvious.
$\diamond$ Exercise 191 Check that $\mathrm{Sp}(2, \mathbb{R})=\operatorname{SL}(2, \mathbb{R})$. (In general, it is not true that $\mathrm{Sp}(2 n, \mathbb{R})=\operatorname{SL}(2 n, \mathbb{R})$.
$\diamond$ Exercise 192 Given $A=\left[\begin{array}{ll}a & b \\ c & d\end{array}\right] \in \mathrm{GL}(2 n, \mathbb{R})$, show that $A \in \operatorname{Sp}(2 n, \mathbb{R})$ if and only if $a^{\top} c$ and $b^{\top} d$ are symmetric and $a^{\top} d-c^{\top} b=I_{n}$.

All matrices of the form

$$
\left[\begin{array}{cc}
I_{n} & 0 \\
t I_{n} & I_{n}
\end{array}\right]
$$

are symplectic. However, the (Frobenius) norm of such a matrix is equal to $\sqrt{2 n+t^{2} n}$, which is unbounded if $t \in \mathbb{R}$. Therefore, $\operatorname{Sp}(2 n, \mathbb{R})$ is not a bounded subset of $\mathbb{R}^{2 n \times 2 n}$ and hence is not compact.

Exercise 193 Consider the skew-symmetric bilinear form on (the vector space) $\mathbb{R}^{2 n}$ defined by

$$
\Omega(x, y):=\sum_{i=1}^{n}\left(x_{i} y_{n+i}-x_{n+i} y_{i}\right)
$$

(the standard symplectic form or the "canonical" symplectic structure). Show that a linear transformation (on $\mathbb{R}^{2 n}$ ) $x \mapsto A x$ preserves the symplectic form $\Omega$ if and only if $A^{\top} \mathbb{J} A=\mathbb{J}$ (i.e., the matrix $A$ is symplectic). Such a structure-preserving transformation is called a symplectic transformation.

The group of all symplectic transformations on $\mathbb{R}^{2 n}$ (equipped with the symplectic form $\Omega$ ) is isomorphic to (the matrix group) $\operatorname{Sp}(2 n, \mathbb{R})$.

Note : The symplectic group is related to classical mechanics. Consider a particle of mass $m$ moving in a potential field $V$. Newton's second law states that the particle moves along a curve $t \mapsto x(t)$ in in Cartesian 3 -space $\mathbb{R}^{3}$ ) in such a way that $m \ddot{x}=-\operatorname{grad} V(x)$. Introduce the conjugate momenta $p_{i}=m \dot{x}_{i}, i=1,2,3$ and the energy (Hamiltonian)

$$
H(x, p):=\frac{1}{2 m} \sum_{i=1}^{3} p_{i}^{2}+V(x)
$$

Then

$$
\frac{\partial H}{\partial x_{i}}=\frac{\partial V}{\partial x_{i}}=-m \ddot{x}_{i}=-\dot{p}_{i} \quad \text { and } \quad \frac{\partial H}{\partial p_{i}}=\frac{1}{m} p_{i}=\dot{x}_{i}
$$

and hence Newton's law $\mathbf{F}=m a$ is equivalent to Hamilton's equations

$$
\dot{x}_{i}=\frac{\partial H}{\partial p_{i}} \quad \text { and } \quad \dot{p}_{i}=-\frac{\partial H}{\partial x_{i}} \quad(i=1,2,3) .
$$

Writing $z=(x, p)$,

$$
\mathbb{J} \cdot \operatorname{grad} H(z)=\left[\begin{array}{cc}
0 & I_{3} \\
-I_{3} & 0
\end{array}\right]\left[\begin{array}{l}
\frac{\partial H}{\partial x} \\
\frac{\partial H}{\partial p}
\end{array}\right]=(\dot{x}, \dot{p})=\dot{z}
$$

so Hamilton equations read $\dot{z}=\mathbb{J} \cdot \operatorname{grad} H(z)$. Now let

$$
F: \mathbb{R}^{3} \times \mathbb{R}^{3} \rightarrow \mathbb{R}^{3} \times \mathbb{R}^{3}
$$

and write $w(t)=F(z(t))$. If $z(t)$ satisfies Hamilton's equations

$$
\dot{z}=\mathbb{J} \cdot \operatorname{grad} H(z)
$$

then $w(t)=F(z(t))$ satisfies $\dot{w}=A^{\top} \dot{z}$, where $A^{\top}=\left[\partial w^{i} / \partial z^{j}\right]$ is the Jacobian matrix of $f$. By the chain rule,

$$
\dot{w}=A^{\top} \mathbb{J} \operatorname{grad}_{z} H(z)=A^{\top} \mathbb{J} A \operatorname{grad}_{w} H(z(w))
$$

Thus, the equations for $w(t)$ have the form of Hamilton's equations with energy $K(w)=H(z(w))$ if and only if $A^{\top} \mathbb{J} A=\mathbb{J}$; that is, if and only if $A$ is symplectic. A nonlinear transformation $F$ is canonical if and only if its Jacobian matrix is symplectic (or, if one prefers, its tangent mapping is a symplectic transformation).

As a special case, consider a (linear transformation) $A \in \operatorname{Sp}(2 n, \mathbb{R})$ and let $w=A z$. Suppose $H$ is quadratic (i.e., of the form $H(z)=\frac{1}{2} z^{\top} B z$ where $B$ is a symmetric matrix). Then $\operatorname{grad} H(z)=B z$ and thus the equations of motion become the linear equations $\dot{z}=\mathbb{J} B z$. Now

$$
\dot{w}=A \dot{z}=A \mathbb{J} B z=\mathbb{J}\left(A^{\top}\right)^{-1} B z=\mathbb{J}\left(A^{\top}\right)^{-1} B A^{-1} A z=\mathbb{J} B^{\prime} w
$$

where $B^{\prime}=\left(A^{\top}\right)^{-1} B A^{-1}$ is symmetric. For the new Hamiltonian we get

$$
\begin{aligned}
H^{\prime}(w) & =\frac{1}{2} w^{\top}\left(A^{\top}\right)^{-1} B A^{-1} w=\frac{1}{2}\left(A^{-1} w\right)^{\top} B A^{-1} w \\
& =H\left(A^{-1} w\right)=H(z)
\end{aligned}
$$

Thus $\operatorname{Sp}(2 n, \mathbb{R})$ is the linear invariance group of classical mechanics.

## The complex general linear group $\mathrm{GL}(n, \mathbb{C})$

Many important matrix groups involve complex matrices. As in the real case,

$$
\mathrm{GL}(n, \mathbb{C}):=\left\{A \in \mathbb{C}^{n \times n} \mid \operatorname{det} A \neq 0\right\}
$$

is an open subset of $\mathbb{C}^{n \times n}$, and hence is not compact. Clearly $G L(n, \mathbb{C})$ is a group under matrix multiplication.

Note : The general linear group $\mathrm{GL}(n, \mathbb{C})$ is connected. This is in contrast with the fact that $\mathrm{GL}(n, \mathbb{R})$ has two components.

## The complex special linear group $\operatorname{SL}(n, \mathbb{C})$

This group is defined by

$$
\mathrm{SL}(n, \mathbb{C}):=\{A \in \mathrm{GL}(n, \mathbb{C}) \mid \operatorname{det} A=1\}
$$

and is treated as in the real case. The matrix group $\operatorname{SL}(n, \mathbb{C})$ is not compact but connected.

The unitary and special unitary groups $\mathrm{U}(n)$ and $\mathrm{SU}(n)$

For $A=\left[a_{i j}\right] \in \mathbb{C}^{n \times n}$,

$$
A^{*}:=\bar{A}^{\top}=\overline{A^{\top}}
$$

is the Hermitian conjugate (i.e., the conjugate transpose) matrix of $A$; thus, $\left(A^{*}\right)_{i j}=\bar{a}_{j i}$. The unitary group is defined as

$$
\mathrm{U}(n):=\left\{A \in \mathrm{GL}(n, \mathbb{C}) \mid A^{*} A=I_{n}\right\} .
$$

$\diamond$ Exercise 194 Verify that $\mathrm{U}(n)$ is a subgroup of the general linear group $\mathrm{GL}(n, \mathbb{C})$.

The unitary condition amounts to $n^{2}$ equations for the $n^{2}$ complex numbers $a_{i j}, i, j=1,2, \ldots, n$

$$
\sum_{k=1}^{n} \bar{a}_{k i} a_{k j}=\delta_{i j} .
$$

By taking real and imaginary parts, these equations actually give $2 n^{2}$ equations in the $2 n^{2}$ real and imaginary parts of the $a_{i j}$ (although there is some redundancy). This means that $\mathrm{U}(n)$ is a closed subset of $\mathbb{C}^{n \times n}=\mathbb{R}^{2 n^{2}}$ and hence of $\mathrm{GL}(n, \mathbb{C})$. Thus $\mathrm{U}(n)$ is a complex matrix group.

Note : The unitary group $U(n)$ is compact and connected.
Let $A \in \mathrm{U}(n)$. From $|\operatorname{det} A|=1$, we see that the determinant function det: $\mathrm{GL}(n, \mathbb{C}) \rightarrow \mathbb{C}$ maps $\mathrm{U}(n)$ onto the unit circle $\mathbb{S}^{1}=\{z \in \mathbb{C}| | z \mid=1\}$.

Note : In the special case $n=1$, a complex linear mapping $\phi: \mathbb{C} \rightarrow \mathbb{C}$ is multiplication by some complex number $z$, and $\phi$ is an isometry if and only if $|z|=1$. In this way, the unitary group $\mathrm{U}(1)$ is identified with the unit circle $\mathbb{S}^{1}$. The group $\mathrm{U}(1)$ is more commonly known as the circle group or the 1 -dimensional torus, and is also denoted by $\mathbb{T}^{1}$.

The dot product on $\mathbb{R}^{n}$ can be extended to $\mathbb{C}^{n}$ by setting (for $x, y \in$ $\mathbb{C}^{n \times 1}$ )

$$
x \bullet y:=x^{*} y=\bar{x}_{1} y_{1}+\bar{x}_{2} y_{2}+\cdots+\bar{x}_{n} y_{n} .
$$

Note : This is not $\mathbb{C}$-linear but satisfies (for $x, y \in \mathbb{C}^{n \times 1}$ and $u, v \in \mathbb{C}$ )

$$
(u x) \bullet(v y)=\bar{u} v(x \bullet y) .
$$

This dot product allows us to define the length (or norm) of a complex vector $x \in \mathbb{C}^{n \times 1}$ by

$$
\|x\|:=\sqrt{x \bullet x}
$$

Then a matrix $A \in \mathbb{C}^{n \times n}$ is unitary if and only if (for $x, y \in \mathbb{C}^{n}$ )

$$
A x \bullet A y=x \bullet y
$$

$\diamond$ Exercise 195 If $G_{i} \leq \mathrm{GL}\left(n_{i}, \mathbb{k}\right), i=1,2$ are matrix groups, show that their (direct) product $G_{1} \times G_{2}$ is also a matrix group (in $\operatorname{GL}\left(n_{1}+n_{2}, \mathbb{k}\right)$ ). Observe, in particular, that the $k$-dimensional torus

$$
\mathbb{T}^{k}:=\mathbb{T}^{1} \times \mathbb{T}^{1} \times \cdots \times \mathbb{T}^{1}
$$

is a matrix group (in $\mathrm{GL}(k, \mathbb{C})$ ). These groups are compact connected Abelian matrix groups. In fact, they are the only matrix groups with these properties.

## The special unitary group

$$
\operatorname{SU}(n):=\{A \in \mathrm{U}(n) \mid \operatorname{det} A=1\}
$$

is a closed subgroup of $\mathrm{U}(n)$ and hence a complex matrix group.
Note : The matrix group $\operatorname{SU}(n)$ is compact and connected. In the special case $n=2, \mathrm{SU}(2)$ is diffeomorphic to the unit sphere $\mathbb{S}^{3}$ in $\mathbb{C}^{2}$ (or $\mathbb{R}^{4}$ ). The group $\mathrm{SU}(2)$ is used in the construction of the gauge group for the Young-Mills equations in elementary particle physics. Also, there is a 2 to 1 surjection (in fact, a surjective submersion)

$$
\pi: \mathrm{SU}(2) \rightarrow \mathrm{SO}(3)
$$

which is of crucial importance in computational mechanics (it is related to the quaternionic representation of rotations in Euclidean 3-space).

## The complex orthogonal groups $\mathrm{O}(n, \mathbb{C})$ and $\mathrm{SO}(n, \mathbb{C})$

Consider the bilinear form on (the vector space) $\mathbb{C}^{n}$ defined by (for $x, y \in$ $\left.\mathbb{C}^{n}\right)$

$$
(x, y):=x_{1} y_{1}+x_{2} y_{2}+\cdots+x_{n} y_{n}
$$

This form is not an inner product because of the lack of complex conjugation in the definition. The set of all complex $n \times n$ matrices which preserve this
form (i.e., such that $(A x, A y)=(x, y)$ for all $\left.x, y \in \mathbb{C}^{n}\right)$ is the complex orthogonal group $\mathrm{O}(n, \mathbb{C})$. Thus

$$
\mathrm{O}(n, \mathbb{C}):=\left\{A \in \mathrm{GL}(n, \mathbb{C}) \mid A^{\top} A=I_{n}\right\} \subseteq \mathrm{GL}(n, \mathbb{C})
$$

It is easy to show that $\mathrm{O}(n, \mathbb{C})$ is a matrix group, and that $\operatorname{det} A= \pm 1$ for all $\mathrm{O}(n, \mathbb{C})$.

Note : The matrix group $\mathrm{O}(n, \mathbb{C})$ is not the same as the unitary group $\mathrm{U}(n)$.

## The complex special orthogonal group

$$
\mathrm{SO}(n, \mathbb{C}):=\{A \in \mathrm{O}(n, \mathbb{C}) \mid \operatorname{det} A=1\}
$$

is also a matrix group.

## The unipotent group $\mathrm{UT}^{u}(n, \mathbb{k})$

A matrix $A=\left[a_{i j}\right] \in \mathbb{k}^{n \times n}$ is upper triangular if all the entries bellow the main diagonal are equal to 0 . Let $\mathrm{UT}(n, \mathbb{k})$ denote the set of all $n \times n$ invertible upper triangular matrices (over $\mathbb{k}$ ). Thus

$$
\mathrm{UT}(n, \mathbb{k}):=\left\{A \in \mathrm{GL}(n, \mathbb{k}) \mid a_{i j}=0 \text { for } i>j\right\} .
$$

$\diamond$ Exercise 196 Show that UT $(n, \mathbb{k})$ is a closed subgroup of the general linear $\operatorname{group} \mathrm{GL}(n, \mathbb{k})$ (and hence a matrix group).

The group $\mathrm{UT}(n, \mathbb{k})$ is called the (real or complex) upper triangular group. This group is not compact.

Note : Likewise, one can define the lower triangular group

$$
\mathrm{LT}(n, \mathbb{k}):=\left\{A \in \mathrm{GL}(n, \mathbb{k}) \mid a_{i j}=0 \text { for } i<j\right\} .
$$

Clearly, $A \in \mathrm{LT}(n, \mathbb{k})$ if and only if $A^{\top} \in \mathrm{UT}(n, \mathbb{k})$. The matrix groups $\mathrm{UT}(n, \mathbb{k})$ and $\mathrm{LT}(n, \mathbb{k})$ are isomorphic and there is no need to distinguish between them.
$\diamond$ Exercise 197 Show that the diagonal group

$$
\mathrm{D}(n, \mathbb{k}):=\left\{A \in \mathrm{GL}(n, \mathbb{k}) \mid a_{i j}=0 \text { for } i \neq j\right\}
$$

is a closed subgroup of $\mathrm{UT}(n, \mathbb{k})$ (and hence a matrix group).

Exercise 198 For $k \leq n$, let $\mathrm{P}(k)$ denote the group of all linear transformations (i.e., invertible linear mappings) on $\mathbb{R}^{n}$ that preserve the subspace $\mathbb{R}^{k}=$ $\mathbb{R}^{k} \times\{0\} \subseteq \mathbb{R}^{n}$. Show that $\mathrm{P}(k)$ is (isomorphic to) the matrix group

$$
\left\{\left.\left[\begin{array}{cc}
A & X \\
0 & B
\end{array}\right] \right\rvert\, A \in \mathrm{GL}(k, \mathbb{R}), B \in \mathrm{GL}(n-k, \mathbb{R}), X \in \mathbb{R}^{k \times(n-k)}\right\}
$$

An upper triangular matrix $A=\left[a_{i j}\right]$ is unipotent if it has all diagonal entries equal to 1. The (real or complex) unipotent group is (the subgroup of $\operatorname{GL}(n, \mathbb{k}))$

$$
\mathrm{UT}^{u}(n, \mathbb{k}):=\left\{A \in \mathrm{GL}(n, \mathbb{k}) \mid a_{i j}=0 \text { for } i>j \text { and } a_{i i}=1\right\}
$$

(see also Exercise 194). It is easy to see that the unipotent group $\mathrm{UT}^{u}(n, \mathbb{k})$ is a closed subgroup of $G L(n, \mathbb{k})$ and hence a matrix group.

Note : $\mathrm{UT}^{u}(n, \mathbb{k})$ is a closed subgroup of $\mathrm{UT}(n, \mathbb{k})$.
For the case

$$
\mathrm{UT}^{u}(2, \mathbb{k})=\left\{\left.\left[\begin{array}{cc}
1 & t \\
0 & 1
\end{array}\right] \in \mathrm{GL}(n, \mathbb{k}) \right\rvert\, t \in \mathbb{k}\right\}
$$

the mapping

$$
\theta: \mathbb{k} \rightarrow \mathrm{UT}^{u}(2, \mathbb{k}), \quad t \mapsto\left[\begin{array}{ll}
1 & t \\
0 & 1
\end{array}\right]
$$

is a continuous group homomorphism which is an isomorphism with continuous inverse. This allows us to view $\mathbb{k}$ as a matrix group.

Note : Given two matrix groups $G$ and $H$, a group homomorphism $\theta: G \rightarrow H$ is a continuous homomorphism if it is continuous and its image $\theta(G) \leq H$ is a closed subset of $H$. For instance,

$$
\theta: \mathrm{UT}^{u}(2, \mathbb{R}) \rightarrow \mathrm{U}(1), \quad\left[\begin{array}{ll}
1 & t \\
0 & 1
\end{array}\right] \mapsto e^{2 \pi t i}
$$

is a continuous homomorphism of matrix groups, but (for $a \in \mathbb{R} \backslash \mathbb{Q}$ )

$$
\theta^{\prime}: G=\left\{\left.\left[\begin{array}{ll}
1 & k \\
0 & 1
\end{array}\right] \in \operatorname{SUT}(2, \mathbb{R}) \right\rvert\, k \in \mathbb{Z}\right\} \rightarrow \mathbf{U}(1), \quad\left[\begin{array}{cc}
1 & k \\
0 & 1
\end{array}\right] \mapsto e^{2 \pi k a i}
$$

is not (since its image is a dense proper subset of $\mathrm{U}(1)$ ). Whenever we have a continuous homomorphism of matrix groups $\theta: G \rightarrow H$ which is a homeomorphism (i.e., a continuous bijection with continuous inverse) we say that $\theta$ is a continuous isomorphism and regard $G$ and $H$ as "identical" (as matrix groups).

The unipotent group $\mathrm{UT}^{u}(3, \mathbb{R})$ is the Heisenberg group

$$
\text { Heis }:=\left\{\left.\left[\begin{array}{ccc}
1 & a & b \\
0 & 1 & c \\
0 & 0 & 1
\end{array}\right] \right\rvert\, a, b, c \in \mathbb{R}\right\}
$$

which is particularly important in quantum physics; the Lie algebra of Heis gives a realization of the Heisenberg commutation relations of quantum mechanics.
$\diamond$ Exercise 199 Verify that the $4 \times 4$ unipotent matrices $A$ of the form

$$
A=\left[\begin{array}{cccc}
1 & a_{2} & a_{3} & a_{4} \\
0 & 1 & a_{1} & \frac{a_{1}^{2}}{2} \\
0 & 0 & 1 & a_{1} \\
0 & 0 & 0 & 1
\end{array}\right]
$$

form a closed subgroup of $\mathrm{UT}^{u}(4, \mathbb{R})$ (and hence a matrix group). Generalize.

Several other matrix groups are of great interest. We describe briefly some of them.

## The general affine group GA $(n, \mathbb{k})$

The general affine group (over $\mathbb{k}$ ) is the group

$$
\mathrm{GA}(n, \mathbb{k}):=\left\{\left.\left[\begin{array}{ll}
1 & 0 \\
c & A
\end{array}\right] \in \mathrm{GL}(n+1, \mathbb{k}) \right\rvert\, c \in \mathbb{k}^{n \times 1} \text { and } A \in \mathrm{GL}(n, \mathbb{k})\right\}
$$

This is clearly a closed subgroup of the general linear group $G L(n+1, \mathbb{k})$ (and hence a matrix group). The general affine group $G A(n, \mathbb{k})$ is not compact. Likewise the case of the general linear group, the matrix group $\mathrm{GA}(n, \mathbb{C})$ is connected but $G A(n, \mathbb{R})$ is not.

NOTE: If we identify the element $x \in \mathbb{k}^{n}$ with $\left[\begin{array}{l}1 \\ x\end{array}\right] \in \mathbb{k}^{(n+1) \times 1}$, then since

$$
\left[\begin{array}{ll}
1 & 0 \\
c & A
\end{array}\right]\left[\begin{array}{l}
1 \\
x
\end{array}\right]=\left[\begin{array}{c}
1 \\
A x+c
\end{array}\right]
$$

we obtain an action of the group $G A(n, \mathbb{k})$ on (the vector space) $\mathbb{k}^{n}$. Transformations on $\mathbb{k}^{n}$ having the form $x \mapsto A x+c$ (with $A$ invertible) are called affine transformations and they preserve lines (i.e., translates of 1-dimensional subspaces of the vector space $\left.\mathbb{k}^{n}\right)$. The associated geometry is affine geometry that has $G A(n, \mathbb{k})$ as its symmetry group.

The (additive group of the) vector space $\mathbb{k}^{n}$ (in fact, $\mathbb{k}^{n \times 1}$ ) can be viewed as (and identified with) the translation subgroup of GA ( $n, \mathbb{k}$ )

$$
\left\{\left.\left[\begin{array}{cc}
1 & 0 \\
c & I_{n}
\end{array}\right] \in \mathrm{GL}(n+1, \mathbb{k}) \right\rvert\, c \in \mathbb{k}^{n \times 1}\right\} \leq \mathrm{GA}(n, \mathbb{k})
$$

and this is a closed subgroup.
The identity component of the (real) general affine group $\mathrm{GA}(n, \mathbb{R})$ is (the matrix group)

$$
\mathrm{GA}^{+}(n, \mathbb{R})=\left\{\left.\left[\begin{array}{ll}
1 & 0 \\
c & A
\end{array}\right] \right\rvert\, c \in \mathbb{K}^{n \times 1} \text { and } A \in \mathrm{GL}^{+}(n, \mathbb{R})\right\}
$$

In particular,

$$
\mathrm{GA}^{+}(1, \mathbb{R})=\left\{\left.\left[\begin{array}{cc}
1 & 0 \\
c & e^{a}
\end{array}\right] \right\rvert\, a, c \in \mathbb{R}\right\}
$$

is a connected matrix group (of "dimension" 2). Its elements are (in fact, can be identified with) transformations on (the real line) $\mathbb{R}$ having the form $x \mapsto b x+c$ (with $b, c \in \mathbb{R}$ and $b>0$ ).

## The Euclidean group $\mathrm{E}(n)$

This is the matrix group

$$
\mathrm{E}(n):=\left\{\left.\left[\begin{array}{ll}
1 & 0 \\
c & A
\end{array}\right] \in \mathrm{GL}(n+1, \mathbb{R}) \right\rvert\, c \in \mathbb{R}^{n \times 1} \text { and } A \in \mathrm{O}(n)\right\} .
$$

The Euclidean group $\mathrm{E}(n)$ is a closed subgroup of the general affine group $\mathrm{GA}(n, \mathbb{R})$ and also is neither compact nor connected. It can be viewed as (and thus identified with) the group of all isometries (i.e., rigid motions) on the Euclidean $n$-space $\mathbb{R}^{n}$.

## The special Euclidean group SE (n)

The special Euclidean group $\operatorname{SE}(n)$ is (the matrix group) defined by

$$
\mathrm{SE}(n):=\left\{\left.\left[\begin{array}{ll}
1 & 0 \\
c & R
\end{array}\right] \in \mathrm{GL}(n+1, \mathbb{R}) \right\rvert\, c \in \mathbb{R}^{n \times 1} \text { and } R \in \mathrm{SO}(n)\right\} .
$$

This group is isomorphic to the group of all orientation-preserving isometries (i.e., proper rigid motions) on the Euclidean n-space $\mathbb{R}^{n}$. It is not compact but connected.

## Some other groups

Several important groups which are not naturally groups of matrices can be viewed as matrix groups. We have seen that the multiplicative groups $\mathbb{R}^{\times}$and $\mathbb{C}^{\times}$(of non-zero real numbers and complex numbers, respectively) are isomorphic to the matrix groups $\mathrm{GL}(1, \mathbb{R})$ and $\mathrm{GL}(1, \mathbb{C})$, respectively. Also, the circle group $\mathbb{S}^{1}$ (of complex numbers with absolute value one) is isomorphic to $\mathrm{U}(1)$. The $n$-torus (the direct product of $n$ copies of $\mathbb{S}^{1}$ )

$$
\mathbb{T}^{n}=\mathbb{S}^{1} \times \cdots \times \mathbb{S}^{1} \leq \mathrm{GL}(n, \mathbb{C})
$$

is isomorphic to the matrix group of $n \times n$ diagonal matrices with complex entries of modulus one. ( $\mathbb{T}^{n}$ can also be realized as the quotient group $\mathbb{R}^{n} / \mathbb{Z}^{n}$ : an element $\left(\theta_{1}, \ldots, \theta_{n}\right) \bmod \mathbb{Z}^{n}$ of $\mathbb{R}^{n} / \mathbb{Z}^{n}$ can be identified with the diagonal matrix $\operatorname{diag}\left(e^{2 \pi i \theta_{1}}, \ldots, e^{2 \pi i \theta_{n}}\right)$.)
Note : If $\theta: G \rightarrow H$ is a continuous homomorphism of matrix groups, then its kernel $\operatorname{ker} \theta \leq G$ is a matrix group. Moreover, the quotient group $G / \operatorname{ker} \theta$ can be identified with the matrix group $\theta(G)$ by the usual quotient isomorphism $\widetilde{\theta}$ : $G / \operatorname{ker} \theta \rightarrow \theta(G)$.

However, it is important to realize that not every normal matrix subgroup $N$ of the matrix group $G$ gives rise to a matrix group $G / N$; there are examples for which
$G / N$ is a Lie group but not a matrix group. (We shall see later that every matrix group is a Lie group.)

Recall that the additive groups $\mathbb{R}$ and $\mathbb{C}$ are isomorphic to the unipotent groups $\mathrm{UT}^{u}(2, \mathbb{R})$ and $\mathrm{UT}^{u}(2, \mathbb{C})$, respectively.
$\diamond$ Exercise 200 Verify that the map

$$
x \in \mathbb{R} \mapsto\left[e^{x}\right] \in \mathrm{GL}^{+}(1, \mathbb{R})
$$

is a continuous isomorphism of matrix groups, and then show that the additive group $\mathbb{R}^{n}$ is isomorphic to the matrix group of all $n \times n$ diagonal matrices with positive entries.
$\diamond$ Exercise 201 Let $\mathbb{Z}^{n} \leq \mathbb{R}^{n}$ be the discrete subgroup of vectors with integer entries and set

$$
\mathrm{GL}(n, \mathbb{Z}):=\left\{A \in \mathrm{GL}(n, \mathbb{R}) \mid A\left(\mathbb{Z}^{n}\right)=\mathbb{Z}^{n}\right\}
$$

Show that $\mathrm{GL}(n, \mathbb{Z})$ is a matrix group. (This matrix group consists of $n \times n$ matrices over (the ring) $\mathbb{Z}$ with determinant $\pm 1$.)

The symmetric group $S_{n}$ of all permutations on $n$ elements may be considered as well as a matrix group. Indeed, we can make $S_{n}$ to act (from the right) on $\mathbb{k}^{n}$ by linear transformations :

$$
\left[\begin{array}{c}
x_{1} \\
x_{2} \\
\vdots \\
x_{n}
\end{array}\right] \cdot \sigma=\left[\begin{array}{c}
x_{\sigma^{-1}(1)} \\
x_{\sigma^{-1}(2)} \\
\vdots \\
x_{\sigma^{-1}(n)}
\end{array}\right] .
$$

Thus (for the standard unit vectors $\left.e_{1}, e_{2}, \ldots, e_{n}\right) e_{i} \cdot \sigma=e_{\sigma(i)}, \quad i=1,2, \ldots, n$.
The matrix $[\sigma]$ of the linear transformation induced by $\sigma \in S_{n}$ (with respect to the standard basis) has all its entries 0 or 1 , with exactly one 1 in each row and column. Such a matrix is usually called a permutation matrix.

Exercise 202 Write down the permutations matrices induces by the elements (permutations) of $S_{3}$.

When $\mathbb{k}=\mathbb{R}$ each of these permutation matrices is orthogonal, while when $\mathbb{k}=\mathbb{C}$ it is unitary. So, for a given $n \geq 1$, the symmetric group $S_{n}$ is (isomorphic to) a closed subgroup of $\mathrm{O}(n)$ or $\mathrm{U}(n)$.

Note : Any finite group is (isomorphic to) a matrix subgroup of some orthogonal group $\mathrm{O}(n)$.

The following table lists some interesting matrix groups, indicates whether or not the group is compact and/or connected, and gives the number of (connected) components.

| Group | Compact ? | Connected ? | Components |
| :---: | :---: | :---: | :---: |
| $\mathrm{GL}(n, \mathbb{C})$ | no | yes | one |
| $\mathrm{SL}(n, \mathbb{C})$ | no | yes | one |
| $\mathrm{GL}(n, \mathbb{R})$ | no | no | two |
| $\mathrm{GL}^{+}(n, \mathbb{R})$ | no | yes | one |
| $\mathrm{SL}(n, \mathbb{R})$ | no | yes | one |
| U (n) | yes | yes | one |
| SU (n) | yes | yes | one |
| $\mathrm{O}(n)$ | yes | no | two |
| SO (n) | yes | yes | one |
| $\mathrm{O}(1, n)$ | no | no | four |
| SO ( $1, n$ ) | no | no | two |
| Lor ( $1, n$ ) | no | yes | one |
| $\mathrm{Sp}(2 n, \mathbb{R})$ | no | yes | one |
| $\mathrm{UT}^{u}(n, \mathbb{k})$ | no | yes | one |
| GA $(n, \mathbb{k})$ | no | no | two |
| $\mathrm{GA}^{+}(n, \mathbb{k})$ | no | yes | one |
| $\mathrm{E}(\mathrm{n})$ | no | no | two |
| SE (n) | no | yes | one |
| $\mathbb{R}^{n}$ | no | yes | one |
| $\mathbb{T}^{n}$ | yes | yes | one |

Note : There are more interesting matrix groups, e.g., the quaternionic matrix groups (in particular, the quaternionic symplectic group $\operatorname{Sp}(n)$ ), associated with the
division algebra $\mathbb{H}$ of quaternions, as well as the spinor groups $\operatorname{Spin}(n)$ and the pinor groups Pin $(n)$, associated with (real) Clifford algebras.

## Complex matrix groups as real matrix groups

Recall that the (complex) vector space $\mathbb{C}$ can be viewed as a real 2-dimensional vector space (with basis $\{1, i\}$, for example).
$\diamond$ Exercise 203 Show that the mapping

$$
\rho: \mathbb{C} \rightarrow \mathbb{R}^{2 \times 2}, \quad z=x+i y \mapsto\left[\begin{array}{cc}
x & -y \\
y & x
\end{array}\right]
$$

is an injective ring homomorphism (i.e., a one-to-one mapping such that, for $z, z^{\prime} \in \mathbb{C}$,

$$
\left.\rho\left(z+z^{\prime}\right)=\rho(z)+\rho\left(z^{\prime}\right) \quad \text { and } \quad \rho\left(z z^{\prime}\right)=\rho(z) \rho\left(z^{\prime}\right) .\right)
$$

We can view $\mathbb{C}$ as a subring of $\mathbb{R}^{2 \times 2}$. In other words, we can identify the complex number $z=x+i y$ with the $2 \times 2$ real matrix $\rho(z)$.

Note : This can also be expressed as

$$
\rho(x+i y)=x I_{2}-y J_{2}, \quad \text { where } \quad J_{2}:=\left[\begin{array}{cc}
0 & 1 \\
-1 & 0
\end{array}\right] .
$$

Also, for $z \in \mathbb{C}$,

$$
\rho(\bar{z})=\rho(z)^{T}
$$

(complex conjugation corresponds to transposition).
More generally, given $Z=\left[z_{r s}\right] \in \mathbb{C}^{n \times n}$ with $z_{r s}=x_{r s}+i y_{r s}$, we can write

$$
Z=X+i Y
$$

where $X=\left[x_{r s}\right], Y=\left[y_{r s}\right] \in \mathbb{R}^{n \times n}$.
Exercise 204 Show that the mapping

$$
\rho_{n}: \mathbb{C}^{n \times n} \rightarrow \mathbb{R}^{2 n \times 2 n}, \quad Z=X+i Y \mapsto\left[\begin{array}{cc}
X & -Y \\
Y & X
\end{array}\right]
$$

is an injective ring homomorphism.

Hence we can identify the complex matrix $Z=X+i Y$ with the $2 n \times 2 n$ real matrix $\rho_{n}(Z)$. Let

$$
\mathbb{J}=\mathbb{J}_{2 n}:=\left[\begin{array}{cc}
0 & I_{n} \\
-I_{n} & 0
\end{array}\right] \in \operatorname{SL}(2 n, \mathbb{R}) .
$$

Then we can write

$$
\rho_{n}(Z)=\rho_{n}(X+i Y)=\left[\begin{array}{cc}
X & 0 \\
0 & X
\end{array}\right]-\left[\begin{array}{cc}
Y & 0 \\
0 & Y
\end{array}\right] \mathbb{J} .
$$

Exercise 205 First verify that

$$
\mathbb{J}^{2}=-I_{2 n} \quad \text { and } \quad \mathbb{J}^{\top}=-\mathbb{J}
$$

and then show that, for $Z \in \mathbb{C}^{n \times n}$,

$$
\rho_{n}(\bar{Z})=\rho_{n}(Z)^{\top} \Longleftrightarrow X=X^{\top} \quad \text { and } \quad Y=Y^{\top} .
$$

We see that $\rho_{n}(\mathrm{GL}(n, \mathbb{C}))$ is a closed subgroup of $\mathrm{GL}(2 n, \mathbb{R})$, so any matrix subgroup $G$ of $\mathrm{GL}(n, \mathbb{C})$ can be viewed as a matrix subgroup of $\mathrm{GL}(2 n, \mathbb{R})$ (by identifying it with its image $\rho_{n}(G)$ under $\left.\rho_{n}\right)$. The following characterizations are sometimes useful :

$$
\begin{aligned}
\rho_{n}\left(\mathbb{C}^{n \times n}\right) & =\left\{A \in \mathbb{R}^{n \times n} \mid A \mathbb{J}=\mathbb{J} A\right\} \\
\rho_{n}(\mathrm{GL}(n, \mathbb{C})) & =\{A \in \mathrm{GL}(2 n, \mathbb{R}) \mid A \mathbb{J}=\mathbb{J} A\} .
\end{aligned}
$$

$\diamond$ Exercise 206 Verify the folowing set of equalities:

$$
\begin{aligned}
\rho_{n}(\mathrm{U}(n)) & =\mathrm{O}(n) \cap \rho_{n}(\mathrm{GL}(n, \mathbb{C})) \\
& =\mathrm{O}(n) \cap \mathrm{Sp}(2 n, \mathbb{R}) \\
& =\rho_{n}(\mathrm{GL}(n, \mathbb{C})) \cap \mathrm{Sp}(2 n, \mathbb{R}) .
\end{aligned}
$$

Note : In a slight abuse of notation, the real symplectic group $\operatorname{Sp}(2 n, \mathbb{R})$ is related to the unitary group $\mathrm{U}(n)$ by

$$
\mathrm{Sp}(2 n, \mathbb{R}) \cap \mathrm{O}(2 n)=\mathrm{U}(n) .
$$

### 4.3 The Exponential Mapping

Let $A \in \mathbb{k}^{n \times n}$ and consider the matrix series

$$
\sum_{k \geq 0} \frac{1}{k!} A^{k}=I_{n}+A+\frac{1}{2!} A^{2}+\frac{1}{3!} A^{3}+\cdots
$$

Note : This matrix series is a series in the complete normed vector space (in fact, algebra) $\left(\mathbb{k}^{n \times n},\|\cdot\|\right)$, where $\|\cdot\|$ is the operator norm (induced by the Euclidean norm on $\left.\mathbb{k}^{n}\right)$. In a complete normed vector space, an absolutely convergent series $\sum_{k \geq 0} a_{k}$ (i.e., such that the series $\sum_{k \geq 0}\left\|a_{k}\right\|$ is convergent) is convergent, and

$$
\left\|\sum_{k=0}^{\infty} a_{k}\right\| \leq \sum_{k=0}^{\infty}\left\|a_{k}\right\| .
$$

(The converse is not true.) Also, every rearrangement of an absolutely convergent series is absolutely convergent, with same sum. Given two absolutely convergent series $\sum_{k \geq 0} a_{k}$ and $\sum_{k \geq 0} b_{k}$ (in a complete normed algebra), their Cauchy product $\sum_{k \geq 0} c_{k}$, where $c_{k}=\sum_{i+j=k} a_{i} b_{j}=a_{0} b_{k}+a_{1} b_{k-1}+\cdots+a_{k} b_{0}$ is also absolutely convergent, and

$$
\sum_{k=0}^{\infty} c_{k}=\left(\sum_{k=0}^{\infty} a_{k}\right)\left(\sum_{k=0}^{\infty} b_{k}\right) .
$$

$\diamond$ Exercise 207 Show that the matrix series $\sum_{k \geq 0} \frac{1}{k!} A^{k}$ is absolutely convergent (and hence convergent).

Let $\sum_{k=0}^{\infty} \frac{1}{k!} A^{k}$ denote the sum of the (absolutely) convergent matrix series $\sum_{k \geq 0} \frac{1}{k!} A^{k}$. We set

$$
e^{A}=\exp (A):=\sum_{k=0}^{\infty} \frac{1}{k!} A^{k}
$$

This matrix is called the matrix exponential of $A$. It follows that

$$
\|\exp (A)\| \leq\left\|I_{n}\right\|+\|A\|+\frac{1}{2!}\|A\|^{2}+\cdots=e^{\|A\|}
$$

and also $\left\|\exp (A)-I_{n}\right\| \leq e^{\|A\|}-1$.

Exercise 208 Show that (for $\lambda, \mu \in \mathbb{k}$ )

$$
\exp ((\lambda+\mu) A)=\exp (\lambda A) \exp (\mu A)
$$

[Hint : These series are absolutely convergent. Think of the Cauchy product.]

It follows that

$$
I_{n}=\exp (O)=\exp ((1+(-1)) A)=\exp (A) \exp (-A)
$$

and hence $\exp (A)$ is invertible with inverse $\exp (-A)$. So $\exp (A) \in G L(n, \mathbb{k})$.

NOTE: The "group property" $\exp ((\lambda+\mu) A)=\exp (\lambda A) \exp (\mu A)$ may be rephrased by saying that, for fixed $A \in \mathbb{k}^{n \times n}$, the mapping $\lambda \mapsto \exp (\lambda A)$ is a (continuous) homomorphism from the additive group of scalars $\mathbb{k}$ into the general linear group $\mathrm{GL}(n, \mathbb{k})$.
4.3.1 Definition. The mapping

$$
\exp : \mathbb{k}^{n \times n} \rightarrow \mathrm{GL}(n, \mathbb{k}), \quad A \mapsto \exp (A)
$$

is called the exponential mapping.
4.3.2 Proposition. If $A, B \in \mathbb{k}^{n \times n}$ commute, then

$$
\exp (A+B)=\exp (A) \exp (B)
$$

Proof : We expand the series and perform a sequence of manipulations that
are legitimate since these series are absolutely convergent :

$$
\begin{aligned}
\exp (A) \exp (B) & =\left(\sum_{r=0}^{\infty} \frac{1}{r!} A^{r}\right)\left(\sum_{s=0}^{\infty} \frac{1}{s!} B^{s}\right) \\
& =\sum_{r, s=0}^{\infty} \frac{1}{r!s!} A^{r} B^{s} \\
& =\sum_{k=0}^{\infty}\left(\sum_{r=0}^{k} \frac{1}{r!(k-r)!} A^{r} B^{k-r}\right) \\
& =\sum_{k=0}^{\infty} \frac{1}{k!}\left(\sum_{r=0}^{k}\binom{k}{r} A^{r} B^{k-r}\right) \\
& =\sum_{k=0}^{\infty} \frac{1}{k!}(A+B)^{k} \\
& =\exp (A+B) .
\end{aligned}
$$

Note : We have made crucial use of the commutativity of $A$ and $B$ in the identity

$$
\sum_{r=0}^{k}\binom{k}{r} A^{r} B^{k-r}=(A+B)^{k}
$$

In particular, for the (commuting) matrices $\lambda A$ and $\mu A$, we reobtain the property $\exp ((\lambda+\mu) A)=\exp (\lambda A) \exp (\mu A)$. It is important to realize that, in fact, the following statements are equivalent (for $A, B \in \mathbb{k}^{n \times n}$ ):
(i) $A B=B A$.
(ii) $\exp (\lambda A) \exp (\mu B)=\exp (\mu B) \exp (\lambda A)$ for all $\lambda, \mu \in \mathbb{k}$.
(iii) $\exp (\lambda A+\mu B)=\exp (\lambda A) \exp (\mu B)$ for all $\lambda, \mu \in \mathbb{k}$.
$\diamond$ Exercise 209 Compute (for $a, b \in \mathbb{R}$ )

$$
\exp \left(\left[\begin{array}{cc}
a & 0 \\
0 & a
\end{array}\right]\right), \quad \exp \left(\left[\begin{array}{cc}
a & -b \\
b & a
\end{array}\right]\right), \quad \exp \left(\left[\begin{array}{ll}
a & b \\
b & a
\end{array}\right]\right), \quad \exp \left(\left[\begin{array}{cc}
a & b \\
0 & a
\end{array}\right]\right)
$$

Note : Every real $2 \times 2$ matrix is conjugate to exactly one of the following types (with $a, b \in \mathbb{R}, b \neq 0$ ) :

- $a\left[\begin{array}{ll}1 & 0 \\ 0 & 1\end{array}\right]$ (scalar).
- $a\left[\begin{array}{ll}1 & 0 \\ 0 & 1\end{array}\right]+b\left[\begin{array}{cc}0 & -1 \\ 1 & 0\end{array}\right]$ (elliptic).
- $a\left[\begin{array}{ll}1 & 0 \\ 0 & 1\end{array}\right]+b\left[\begin{array}{ll}0 & 1 \\ 1 & 0\end{array}\right]$ (hyperbolic).
- $a\left[\begin{array}{ll}1 & 0 \\ 0 & 1\end{array}\right]+b\left[\begin{array}{ll}0 & 1 \\ 0 & 0\end{array}\right]$ (parabolic).


## Hermitian and skew-Hermitian matrices.

$\diamond$ Exercise 210
(a) Show that if $A \in \mathbb{R}^{n \times n}$ is skew-symmetric, then $\exp (A)$ is orthogonal.
(b) Show that if $A \in \mathbb{C}^{n \times n}$ is skew-Hermitian, then $\exp (A)$ is unitary.
$\diamond$ Exercise 211 Let $A \in \mathbb{k}^{n \times n}$ and $B \in G L(n, \mathbb{k})$. Show that

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

Deduce that if $B^{-1} A B=\operatorname{diag}\left(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{n}\right)$, then

$$
\exp (A)=B \operatorname{diag}\left(e^{\lambda_{1}}, e^{\lambda_{2}}, \ldots, e^{\lambda_{n}}\right) B^{-1}
$$

$\diamond$ Exercise 212 A matrix $A \in \mathbb{k}^{n \times n}$ is nilpotent if $A^{k}=O$ for some $k \geq 1$.
(a) Prove that a nilpotent matrix is singular.
(b) Prove that a strictly upper triangular matrix $A=\left[a_{i j}\right]$ (i.e. with $a_{i j}=0$ whenever $i \geq j$ ) is nilpotent.
(c) Find two nilpotent matrices whose product is not nilpotent.

Exercise 213 Suppose that $A \in \mathbb{k}^{n \times n}$ and $\|A\|<1$.
(a) Show that the matrix series

$$
\sum_{k \geq 0} A^{k}=I_{n}+A+A^{2}+A^{3}+\cdots
$$

converges (in $\mathbb{k}^{n \times n}$ ).
(b) Show that the matrix $I_{n}-A$ is invertible and find a formula for ( $I_{n}-$ $A)^{-1}$.
(c) If $A$ is nilpotent, determine $\left(I_{n}-A\right)^{-1}$ and $\exp (A)$.
$\diamond$ Exercise 214 Show (for $\lambda \in \mathbb{R}$ )

$$
\exp \left(\left[\begin{array}{ccccc}
\lambda & 1 & 0 & \ldots & 0 \\
0 & \lambda & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \ldots & \lambda
\end{array}\right]\right)=\left[\begin{array}{ccccc}
e^{\lambda} & e^{\lambda} & \frac{1}{2!} e^{\lambda} & \ldots & \frac{1}{(n-1)!} e^{\lambda} \\
0 & e^{\lambda} & e^{\lambda} & \ldots & \frac{1}{(n-2)!} e^{\lambda} \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \ldots & e^{\lambda}
\end{array}\right]
$$

Note : When the matrix $A \in \mathbb{k}^{n \times n}$ is diagonalizable over $\mathbb{C}$ (i.e., $A=$ $C \operatorname{diag}\left(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{n}\right) C^{-1}$ for some $C \in \mathrm{GL}(n, \mathbb{C})$ ), we have

$$
\exp (A)=C \operatorname{diag}\left(e^{\lambda_{1}}, e^{\lambda_{2}}, \ldots, e^{\lambda_{n}}\right) C^{-1}
$$

This means that the problem of calculating the exponential of a diagonalizable matrix is solved once an explicit diagonalization is found. Many important types of matrices are indeed diagonalizable (over $\mathbb{C}$ ), including skew-symmetric, skew-Hermitian, orthogonal, and unitary matrices. However, there are also many non-diagonalizable matrices. If $A^{k}=O$ for some positive integer $k$, then $A^{\ell}=O$ for all $\ell \geq k$. In this case the matrix series which defines $\exp (A)$ terminates after the first $k$ terms, and so can be computed explicitly. A general matrix $A$ may be neither nilpotent nor diagonalizable. This situation is best discussed in terms of the Jordan canonical form.

For $\lambda \in \mathbb{C}$ and $r \geq 1$, we have the Jordan block matrix

$$
J(\lambda, r):=\left[\begin{array}{cccccc}
\lambda & 1 & 0 & \ldots & 0 & 0 \\
0 & \lambda & 1 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \ldots & \lambda & 1 \\
0 & 0 & 0 & \ldots & 0 & \lambda
\end{array}\right] \in \mathbb{C}^{r \times r}
$$

The characteristic polynomial of $J(\lambda, r)$ is

$$
\operatorname{char}_{J(\lambda, r)}(s):=\operatorname{det}\left(s I_{r}-J(\lambda, r)\right)=(s-\lambda)^{r}
$$

and by the Cayley-Hamilton Theorem, $\left(J(\lambda, r)-\lambda I_{r}\right)^{r}=O$, which implies that $\left(J(\lambda, r)-\lambda I_{r}\right)^{r-1} \neq O$ (and hence $\left.\operatorname{char}_{J(\lambda, r)}(s)=\min _{J(\lambda, r)}(s) \in \mathbb{C}[s]\right)$. The main result on Jordan form is the following : Given $A \in \mathbb{C}^{n \times n}$, there exists a matrix
$P \in \mathrm{GL}(n, \mathbb{C})$ such that

$$
P^{-1} A P=\left[\begin{array}{cccc}
J\left(\lambda_{1}, r_{1}\right) & O & \cdots & O \\
O & J\left(\lambda_{2}, r_{2}\right) & \ldots & O \\
\vdots & \vdots & & \vdots \\
O & O & \ldots & J\left(\lambda_{m}, r_{m}\right)
\end{array}\right] \in \mathbb{C}^{n \times n}
$$

This form is unique except for the order in which the Jordan blocks $J\left(\lambda_{i}, r_{i}\right) \in \mathbb{C}^{r_{i} \times r_{i}}$ occur. (The elements $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{m}$ are the eigenvalues of $A$ and in fact $\operatorname{char}_{A}(s)=$ $\left.\left(s-\lambda_{1}\right)^{r_{1}}\left(s-\lambda_{2}\right)^{r_{2}} \cdots\left(s-\lambda_{m}\right)^{r_{m}}.\right)$

Using the Jordan canonical form we can see that every matrix $A \in \mathbb{C}^{n \times n}$ can be written as $A=S+N$, where $S$ is diagonalizable (over $\mathbb{C}$ ), $N$ is nilpotent, and $S N=N S$.
$\diamond$ Exercise 215 Let $A \in \mathbb{k}^{n \times n}$.
(a) Prove that $A$ is nilpotent if and only if all its eigenvalues are equal to zero.
(b) The matrix $A$ is called unipotent if $I_{n}-A$ is nilpotent (i.e. $\left(I_{n}-A\right)^{k}=$ $O$ for some $k \geq 1$ ). Prove that $A$ is unipotent if and only if all its eigenvalues are equal to 1 .
(c) If $A$ is a strictly upper triangular matrix, show that $\exp (A)$ is unipotent.

Exercise 216 Compute

$$
\exp \left(\left[\begin{array}{lll}
\lambda & a & b \\
0 & \lambda & c \\
0 & 0 & \lambda
\end{array}\right]\right)
$$

Exercise 217 The power series

$$
\sum_{k \geq 1}(-1)^{k+1} \frac{(z-1)^{k}}{k}=z-1-\frac{(z-1)^{2}}{2}+\frac{(z-1)^{3}}{3}-\frac{(z-1)^{4}}{4}+\cdots, \quad z \in \mathbb{C}
$$

has radius of convergence 1 and hence defines a complex analytic function

$$
z \mapsto \log z:=\sum_{k=1}^{\infty}(-1)^{k+1} \frac{(z-1)^{k}}{k}
$$

on the set $\{z||z-1|<1\}$. (This function coincides with the usual logarithm for real $z$ on the interval $(0,2)$.) Show that
(a) For all $z$ with $|z-1|<1$,

$$
e^{\log z}=z
$$

(b) For all $w$ with $|w|<\ln 2,\left|e^{w}-1\right|<1$ and

$$
\log \left(e^{w}\right)=w
$$

Let $A \in \mathbb{K}^{n \times n}$. The matrix series

$$
\sum_{k \geq 1} \frac{(-1)^{k+1}}{k} A^{k}=A-\frac{1}{2} A+\frac{1}{3} A^{3}-\frac{1}{4} A^{4}+\cdots
$$

converges (absolutely) for $\|A\|<1$. We define the logarithm mapping

$$
\log : \mathcal{B}_{\mathbb{k}^{n \times n}}\left(I_{n}, 1\right) \rightarrow \mathbb{k}^{n \times n}, \quad A \mapsto \log (A):=\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\left(A-I_{n}\right)^{k}
$$

(The notation $\mathcal{B}_{\mathbb{k}^{n \times n}}(A, \rho)$ stands for the open ball of radius $\rho$ around $A$ in the metric space $\mathbb{k}^{n \times n}$; that is,

$$
\left.\mathcal{B}_{\mathbb{K}^{n \times n}}(A, \rho):=\left\{A^{\prime} \in \mathbb{k}^{n \times n} \mid\left\|A^{\prime}-A\right\|<\rho\right\} .\right)
$$

Note : Defining a logarithm for matrices turns out to be at least as difficult as defining a logarithm for complex numbers, and so we cannot hope to define the matrix logarithm for all matrices, or even for all invertible matrices. We content ourselves with defining the logarithm in a neighborhood of the identity matrix. The logarithm mapping is continuous (on the set of all $n \times n$ matrices $A$ with $\left\|A-I_{n}\right\|<1$ ) and $\log (A)$ is real if $A$ is real.
$\diamond$ Exercise 218 Show that
(a) For all $A$ with $\left\|A-I_{n}\right\|<1$,

$$
\exp (\log (A))=A
$$

(b) For all $B$ with $\|B\|<\ln 2,\left\|\exp (B)-I_{n}\right\|<1$ and

$$
\log (\exp (B))=B
$$

The exponential and logarithm mappings

$$
\exp : \mathbb{k}^{n \times n} \rightarrow \mathrm{GL}(n, \mathbb{k}), \quad \text { and } \quad \log : \mathcal{B}_{\mathbb{k}^{n \times n}}\left(I_{n}, 1\right) \rightarrow \mathbb{k}^{n \times n}
$$

are continuous (in fact, infinitely differentiable). Indeed, since any power $A^{k}$ is a continuous mapping of $A$, the sequence of partial sums $\left(\sum_{k=0}^{r} \frac{1}{k!} A^{k}\right)_{r \geq 0}$ consists of continuous mappings. But (it is easy to see that) the matrix series (defining the exponential matrix) converges uniformly on each set of the form $\{A \mid\|A\| \leq \rho\}$, and so the sum (i.e., the limit of its sequence of partial sums) is again continuous. (A similar argument works in the case of the logarithm mapping.)

By continuity (of the exponential mapping at the origin $O$ ), there is a number $\delta>0$ such that

$$
\mathcal{B}_{\mathbb{k}^{n \times n}}(O, \delta) \subseteq \exp ^{-1}\left(\mathcal{B}_{\mathrm{GL}(n, \mathbf{k})}\left(I_{n}, 1\right)\right) .
$$

In fact we can actually take $\delta=\ln 2$ since

$$
\exp \left(\mathcal{B}_{\mathbb{k}^{n \times n}}(O, \delta)\right) \subseteq \mathcal{B}_{\mathbb{k}^{n \times n}}\left(I_{n}, e^{\delta}-1\right) .
$$

Hence we have the following result
4.3.3 Proposition. The exponential mapping exp is injective when restricted to the open subset $\mathcal{B}_{\mathbb{k}^{n \times n}}(O, \ln 2)$. (Hence it is locally a diffeomorphism at the origin $O$ with local inverse log.)

Let $A \in \mathbb{k}^{n \times n}$. For every $t \in \mathbb{R}$, the matrix series $\sum_{k \geq 0} \frac{t^{k}}{k!} A^{k}$ is (absolutely) convergent and we have

$$
\sum_{k=0}^{\infty} \frac{t^{k}}{k!} A^{k}=\sum_{k=0}^{\infty} \frac{1}{k!}(t A)^{k}=\exp (t A) .
$$

So the mapping

$$
\alpha: \mathbb{R} \rightarrow \mathbb{k}^{n \times n}, \quad t \mapsto \exp (t A)
$$

is defined and differentiable with

$$
\dot{\alpha}(t)=\sum_{k=1}^{\infty} \frac{t^{k-1}}{(k-1)!} A^{k}=\exp (t A) A=A \exp (t A) .
$$

Note : This mapping can be viewed as a curve in $\mathbb{k}^{n \times n}$. The curve is in fact smooth (i.e., infinitely differentiable) and satisfies the differential equation (in matrices) $\dot{\alpha}(t)=\alpha(t) A$ with initial condition $\alpha(0)=I_{n}$. Also (for $t, s \in \mathbb{R}$ ),

$$
\alpha(t+s)=\alpha(t) \alpha(s) .
$$

In particular, this shows that $\alpha(t)$ is always invertible with $\alpha(t)^{-1}=\alpha(-t)$.
$\diamond$ Exercise 219 Let $A, C \in \mathbb{k}^{n \times n}$. Show that the differential equation (in matrices) $\dot{\alpha}=\alpha A$ has a unique differentiable solution $\alpha: \mathbb{R} \rightarrow \mathbb{k}^{n \times n}$ for which $\alpha(0)=C$. (This solution is $\alpha(t)=C \exp (t A)$.) Furthermore, if $C$ is invertible, then so is $\alpha(t)$ for $t \in \mathbb{R}$, hence $\alpha: \mathbb{R} \rightarrow \operatorname{GL}(n, \mathbb{k})$.
$\diamond$ Exercise 220 Let $A \in \mathbb{k}^{n \times n}$. Show that the functional equation (in matrices) $\alpha(t+s)=\alpha(t) \alpha(s)$ has a unique differentiable solution $\alpha: \mathbb{R} \rightarrow \mathbb{k}^{n \times n}$ for which $\alpha(0)=I_{n}$ and $\dot{\alpha}(0)=A$. (This solution is $\alpha(t)=\exp (t A)$.)
$\diamond$ Exercise 221 If $A, B \in \mathbb{k}^{n \times n}$ commute, show that

$$
\left.\frac{d}{d t} \exp (A+t B)\right|_{t=0}=\exp (A) B=B \exp (A)
$$

(This is a formula for the derivative of the exponential mapping exp at an arbitrary $A$, evaluated only at those $B$ such that $A B=B A$. The general situation is more complicated.)

### 4.4 Lie Algebras for Matrix Groups

## One-parameter subgroups

Let $G \leq G L(n, \mathbb{k})$ be a matrix group and let $I$ denote the identity matrix.
Note : The matrix $I$ is the neutral element of the group $G$. When $\mathbb{k}=\mathbb{R}$, then $I=I_{n}$ whereas when $\mathbb{k}=\mathbb{C}$ and $G \leq \mathrm{GL}(2 n, \mathbb{R})$, then $I=I_{2 n}=\left[\begin{array}{cc}I_{n} & 0 \\ 0 & I_{n}\end{array}\right]$.
4.4.1 Definition. A one-parameter subgroup of $G$ is a continuous mapping

$$
\gamma: \mathbb{R} \rightarrow G
$$

which is differentiable at 0 and satisfies (for $t, s \in \mathbb{R}$ )

$$
\gamma(s+t)=\gamma(s) \gamma(t)
$$

We refer to the last condition as the homomorphism property.
Note : Recall that $\mathbb{R}$ and $G$ can be viewed as matrix groups (isomorphic to $\mathrm{UT}^{u}(2, \mathbb{R})$ and to a subgroup of either $\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(2 n, \mathbb{R})$, respectively). Hence, $\gamma$ is a continuous homomorphism of matrix groups.

It suffices to know $\gamma$ on some open neighborhood $(-\varepsilon, \varepsilon)$ of 0 in $\mathbb{R}$. Indeed, let $t \in \mathbb{R}$. Then for some (large enough) natural number $m, \frac{t}{m} \in(-\varepsilon, \varepsilon)$. Hence

$$
\gamma\left(\frac{t}{m}\right),\left(\gamma\left(\frac{t}{m}\right)\right)^{m} \in G
$$

$\diamond$ Exercise 222 Show that (for $m, n \in \mathbb{N}$ such that $\frac{t}{m}, \frac{t}{n} \in(-\varepsilon, \varepsilon)$ )

$$
\left(\gamma\left(\frac{t}{n}\right)\right)^{n}=\left(\gamma\left(\frac{t}{m}\right)\right)^{m}
$$

The element $\left(\gamma\left(\frac{t}{m}\right)\right)^{m} \in G$ is well defined (for every $t \in \mathbb{R}$ ), and so

$$
\gamma(t)=\gamma\left(\frac{t}{m}+\frac{t}{m}+\cdots+\frac{t}{m}\right)=\left(\gamma\left(\frac{t}{m}\right)\right)^{m}
$$

Note : A one-parameter subgroup $\gamma: \mathbb{R} \rightarrow G$ can be viewed as a collection $(\gamma(t))_{t \in \mathbb{R}}$ of linear transformations on $\mathbb{k}^{n}$ such that (for $t, s \in \mathbb{R}$ )

- $\quad \gamma(0)=i d_{\mathbb{k}^{n}}(=I)$.
- $\quad \gamma(s+t)=\gamma(s) \gamma(t)$.
- $\quad \gamma(t) \in G$ depends continuously on $t$.

Moreover, the curve $\gamma: \mathbb{R} \rightarrow G$ in $G \subseteq \mathbb{k}^{n \times n}$ has a tangent vector $\dot{\gamma}(0)$ (at $\gamma(0)=I$ ).
4.4.2 Proposition. Let $\gamma: \mathbb{R} \rightarrow G$ be a one-parameter subgroup of $G$. Then $\gamma$ is differentiable at every $t \in \mathbb{R}$ and

$$
\dot{\gamma}(t)=\dot{\gamma}(0) \gamma(t)=\gamma(t) \dot{\gamma}(0)
$$

Proof : We have (for $t, h \in \mathbb{R}$ )

$$
\begin{aligned}
\dot{\gamma}(t) & =\lim _{h \rightarrow 0} \frac{1}{h}(\gamma(t+h)-\gamma(t)) \\
& =\lim _{h \rightarrow 0} \frac{1}{h}(\gamma(h) \gamma(t)-\gamma(t)) \\
& =\left(\lim _{h \rightarrow 0} \frac{1}{h}(\gamma(h)-I)\right) \gamma(t) \\
& =\dot{\gamma}(0) \gamma(t)
\end{aligned}
$$

and similarly

$$
\dot{\gamma}(t)=\gamma(t) \dot{\gamma}(0) .
$$

We can now determine the form of all one-parameter subgroups of $G$.
4.4.3 Theorem. Let $\gamma: \mathbb{R} \rightarrow G$ be a one-parameter subgroup of $G$. Then it has the form

$$
\gamma(t)=\exp (t A)
$$

for some $A \in \mathbb{k}^{n \times n}$.
Proof : Let $A=\dot{\gamma}(0)$. This means that $\gamma$ satisfies (the differential equation)

$$
\dot{\gamma}(t)=A \gamma(t)
$$

and is subject to (the initial condition)

$$
\gamma(0)=I .
$$

This initial value problem (IVP) has the unique solution $\gamma(t)=\exp (t A)$.
We cannot yet reverse this process and decide for which $A \in \mathbb{k}^{n \times n}$ the one-parameter subgroup

$$
\gamma: \mathbb{R} \rightarrow \mathrm{GL}(n, \mathbb{k}), \quad t \mapsto \exp (t A)
$$

actually takes values in $G$. The answer involves the Lie algebra of $G$.
Note : We have a curious phenomenon in the fact that although the definition of a one-parameter group only involves first order differentiability, the general form
$\exp (t A)$ is always infinitely differentiable (and indeed analytic) as a function of $t$. This is an important characteristic of much of the Lie theory, namely that conditions of first order diferentiability (and even continuity) often lead to much stronger conditions.

## Lie algebras

Let $G \leq \mathrm{GL}(n, \mathbb{k})$ be a matrix group. Recall that $\mathbb{k}^{n \times n}$ may be considered to be some Euclidean space $\mathbb{R}^{m}$.
4.4.4 Definition. A curve in $G$ is a differentiable mapping

$$
\gamma:(a, b) \subseteq \mathbb{R} \rightarrow \mathbb{k}^{n \times n}
$$

such that (for $t \in(a, b)$ )

$$
\gamma(t) \in G
$$

The derivative

$$
\dot{\gamma}(t):=\lim _{h \rightarrow 0} \frac{1}{h}(\gamma(t+h)-\gamma(t)) \in \mathbb{k}^{n \times n}
$$

is called the tangent vector to $\gamma$ at $\gamma(t)$. We will usually assume that $a<0<b$.

Exercise 223 Given two curves $\gamma, \sigma:(a, b) \rightarrow G$, we define a new curve, the product curve, by

$$
(\gamma \sigma)(t):=\gamma(t) \sigma(t)
$$

Show that (for $t \in(a, b)$ )

$$
(\gamma \sigma)^{\cdot}(t)=\gamma(t) \dot{\sigma}(t)+\dot{\gamma}(t) \sigma(t)
$$

$\diamond$ Exercise 224
(a) Let $\gamma:(-1,1) \rightarrow \mathbb{R}^{3 \times 3}$ be given by

$$
\gamma(t):=\left[\begin{array}{ccc}
1 & 0 & 0 \\
0 & \cos t & \sin t \\
0 & -\sin t & \cos t
\end{array}\right]
$$

Show that $\gamma$ is a curve in $\mathbf{S O}(3)$ and find $\dot{\gamma}(0)$. Show that

$$
\left(\gamma^{2}\right)^{\cdot}(0)=2 \dot{\gamma}(0)
$$

(b) Let $\sigma:(-1,1) \rightarrow \mathbb{R}^{3 \times 3}$ be given by

$$
\sigma(t):=\left[\begin{array}{ccc}
0 & 0 & 0 \\
0 & \cos t & \sin t \\
0 & -\sin t & \cos t
\end{array}\right]
$$

Calculate $\dot{\sigma}(0)$. Write the matrix $\gamma(t) \sigma(t)$ and verify that

$$
(\gamma \sigma)^{\cdot}(0)=\dot{\gamma}(0)+\dot{\sigma}(0)
$$

$\diamond$ Exercise 225 Let $\alpha:(-1,1) \rightarrow \mathbb{C}^{n \times n}$ be given by

$$
\alpha(t):=\left[\begin{array}{ccc}
e^{i \pi t} & 0 & \\
0 & e^{i \frac{\pi t}{2}} & 0 \\
0 & 0 & e^{i \frac{\pi t}{2}}
\end{array}\right] .
$$

Show that $\alpha$ is a curve in $\mathrm{U}(3)$. Calculate $\dot{\alpha}(0)$.
4.4.5 Definition. The tangent space to (the matrix group) $G$ at $A \in G$ is the set

$$
T_{A} G:=\left\{\dot{\gamma}(0) \in \mathbb{k}^{n \times n} \mid \gamma \text { is a curve in } G \text { with } \gamma(0)=A\right\}
$$

4.4.6 Proposition. The set $T_{A} G$ is a real vector subspace of $\mathbb{k}^{n \times n}$.

Proof : Let $\alpha, \beta:(a, b) \rightarrow \mathbb{k}^{n \times n}$ be two curves in $G$ through $A$ (i.e., $\alpha(0)=\beta(0)=A)$. Then

$$
\gamma:(a, b) \rightarrow \mathbb{k}^{n \times n}, \quad t \mapsto \alpha(t) A^{-1} \beta(t)
$$

is also a curve in $G$ with $\gamma(0)=A$. We have

$$
\dot{\gamma}(t)=\dot{\alpha}(t) A^{-1} \beta(t)+\alpha(t) A^{-1} \dot{\beta}(t)
$$

and hence

$$
\dot{\gamma}(0)=\dot{\alpha}(0) A^{-1} \beta(0)+\alpha(0) A^{-1} \dot{\beta}(0)=\dot{\alpha}(0)+\dot{\beta}(0)
$$

which shows that $T_{A} G$ is closed under (vector) addition.
Similarly, if $\lambda \in \mathbb{R}$ and $\alpha:(a, b) \rightarrow \mathbb{k}^{n \times n}$ is a curve in $G$ with $\alpha(0)=A$, then

$$
\eta:(a, b) \rightarrow \mathbb{k}^{n \times n}, \quad t \mapsto \alpha(\lambda t)
$$

is another such curve. Since

$$
\dot{\eta}(0)=\lambda \dot{\alpha}(0)
$$

we see that $T_{A} G$ is closed under (real) scalar multiplication. So $T_{A} G$ is a (real) vector subspace of $\mathbb{k}^{n \times n}$.

NOTE : Since the vector space $\mathbb{k}^{n \times n}$ is finite dimensional, so is (the tangent space) $T_{A} G$.
4.4.7 Definition. If $G \leq G L(n, \mathbb{k})$ is a matrix group, its dimension is the dimension of the (real) vector space $T_{I} G$ ( $I$ is the identity matrix). So

$$
\operatorname{dim} G:=\operatorname{dim}_{\mathbb{R}} T_{I} G .
$$

Note : If the matrix group $G$ is complex, then its complex dimension is

$$
\operatorname{dim}_{\mathbb{C}} G:=\operatorname{dim}_{\mathbb{C}} T_{I} G .
$$

$\diamond$ Exercise 226 Show that the matrix group U(1) has dimension 1.
Note : The only connected matrix groups (up to isomorphism) of dimension 1 are $\mathbb{T}^{1}=U(1)$ and $\mathbb{R}$, and of dimension 2 are $\mathbb{R}^{2}, \mathbb{T}^{1} \times \mathbb{R}, \mathbb{T}^{2}$, and $\mathrm{GA}^{+}(1, \mathbb{R})$.
4.4.8 Example. The real general linear group $\mathrm{GL}(n, \mathbb{R})$ has dimension $n^{2}$.

The determinant function det: $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ is continuous and $\operatorname{det}(I)=1$. So there is some $\epsilon$-ball about $I$ in $\mathbb{R}^{n \times n}$ such that for each $A$ in this ball $\operatorname{det}(A) \neq 0$ (i.e., $A \in \mathrm{GL}(n, \mathbb{R}))$. If $B \in \mathbb{R}^{n \times n}$, then define a curve $\sigma$ in $\mathbb{R}^{n \times n}$ by

$$
\sigma(t):=t B+I
$$

Then $\sigma(0)=I$ and $\dot{\sigma}(0)=B$, and (for small $t$ ) $\sigma(t) \in \mathrm{GL}(n, \mathbb{R})$. Hence the tangent space $T_{I} \mathrm{GL}(n, \mathbb{R})$ is all of $\mathbb{R}^{n \times n}$ which has dimension $n^{2}$. So

$$
\operatorname{dim} \mathrm{GL}(n, \mathbb{R})=n^{2}
$$

Exercise 227 Show that the dimension of the complex general linear group $\mathrm{GL}(n, \mathbb{C})$ is $2 n^{2}$.
4.4.9 Proposition. Let Sk-sym ( $n$ ) denote the set of all skew-symmetric matrices in $\mathbb{R}^{n \times n}$. Then Sk-sym $(n)$ is a linear subspace of $\mathbb{R}^{n \times n}$ and its dimension is $\frac{n(n-1)}{2}$.

Proof : If $A, B \in \operatorname{Sk}$-sym $(n)$, then

$$
(A+B)^{\top}+(A+B)=A^{\top}+A+B^{\top}+B=0
$$

so that Sk-sym $(n)$ is closed under (vector) addition.
It is also closed under scalar multiplication, for if $A \in \operatorname{Sk-sym}(n)$ and $\lambda \in \mathbb{R}$, then $(\lambda A)^{\top}=\lambda A^{\top}$ so that

$$
(\lambda A)^{\top}+\lambda A=\lambda\left(A^{\top}+A\right)=0
$$

To check the dimension of $\operatorname{Sk}$-sym $(n)$ we construct a basis. Let $E_{i j}$ denote the matrix whose entries are all zero except the $i j$-entry, which is 1 , and the $j i$-entry, which is -1 . If we define these $E_{i j}$ only for $i<j$, we can see that they form a basis for $\operatorname{Sk}$-sym ( $n$ ).

It is easy to compute that there are

$$
(n-1)+(n-2)+\cdots+2+1=\frac{n(n-1)}{2}
$$

of them.
$\diamond$ Exercise 228 Show that if $\sigma$ is a curve through the identity (i.e., $\sigma(0)=I$ ) in the orthogonal group $\mathrm{O}(n)$, then $\dot{\sigma}(0)$ is skew-symmetric.

Note : It follows that $\operatorname{dim} \mathrm{O}(n) \leq \frac{n(n-1)}{2}$. Later we will show that this evaluation is an equality.
$\diamond$ Exercise 229 A matrix $A \in \mathbb{C}^{n \times n}$ is called skew-Hermitian if $A^{*}+A=0$.
(a) Show that the diagonal terms of a skew-Hermitian matrix are purely imaginary and hence deduce that the set Sk-Herm ( $n$ ) of all skew-Hermitian matrices in $\mathbb{C}^{n \times n}$ is not a vector space over $\mathbb{C}$.
(b) Prove that Sk - $\operatorname{Herm}(n)$ is a real vector space of dimension

$$
n+2 \frac{n(n-1)}{2}=n^{2} .
$$

(c) If $\sigma$ is a curve through the identity in $\mathrm{U}(n)$, show that $\dot{\sigma}(0)$ is skewHermitian an hence

$$
\operatorname{dim} \mathrm{U}(n) \leq n^{2}
$$

We will adopt the notation $\mathfrak{g}:=T_{I} G$ for this real vector subspace of $\mathbb{k}^{n \times n}$. In fact, $\mathfrak{g}$ has a more interesting algebraic structure, namely that of a Lie algebra.

Note : It is customary to use lower case Gothic (Fraktur) characters (such as $\mathfrak{a}, \mathfrak{g}$ and $\mathfrak{h}$ ) to refer to Lie algebras.
4.4.10 Definition. A (real)Lie algebra $\mathfrak{a}$ is a real vector space equipped with a product

$$
[\cdot, \cdot]: \mathfrak{a} \times \mathfrak{a} \rightarrow \mathfrak{a}, \quad(x, y) \mapsto[x, y]
$$

such that (for $\lambda, \mu \in \mathbb{R}$ and $x, y, z \in \mathfrak{a}$ )
(LA1) $\quad[x, y]=-[y, x]$.
(LA2) $\quad[\lambda x+\mu y, z]=\lambda[x, z]+\mu[y, z]$.
$(\mathrm{LA} 3) \quad[x,[y, z]]+[y,[z, x]]+[z,[x, y]]=0$.
The product $[\cdot, \cdot]$ is called the Lie bracket of the Lie algebra $\mathfrak{a}$.
Note : (1) Condition (LA3) is called the Jacobi identity. So the Lie bracket $[\cdot, \cdot]$ of (the Lie algebra) $\mathfrak{a}$ is a skew-symmetric bilinear mapping (on $\mathfrak{a}$ ) which satisfies the Jacobi identity. Hence Lie algebras are nonassociative algebras. The Lie bracket plays for Lie algebras the same role that the associative law plays for associative algebras.
(2) While we can define complex Lie algebras (or, more generally, Lie algebras over any field), we shall only consider Lie algebras over the real field $\mathbb{R}$.
4.4.11 EXAMPLE. Let $\mathfrak{a}=\mathbb{R}^{n}$ and set (for all $x, y \in \mathbb{R}^{n}$ )

$$
[x, y]:=0
$$

The trivial product is a skew-symmetric bilinear multiplication (on $\mathbb{R}^{n}$ ) which satisfies the Jacobi identity and hence is a Lie bracket. $\mathbb{R}^{n}$ equipped with this product (Lie bracket) is a Lie algebra. Such a Lie algebra is called an Abelian Lie algebra.

Exercise 230 Show that the only Lie algebra structure on (the vector space) $\mathbb{R}$ is the trivial one.
4.4.12 EXAMPLE. Let $\mathfrak{a}=\mathbb{R}^{3}$ and set (for $x, y \in \mathbb{R}^{3}$ )

$$
[x, y]:=x \times y \quad \text { (the cross product }) .
$$

For the standard unit vectors $e_{1}, e_{2}, e_{3}$ we have

$$
\left[e_{1}, e_{2}\right]=-\left[e_{2}, e_{1}\right]=e_{3}, \quad\left[e_{2}, e_{3}\right]=-\left[e_{3}, e_{2}\right]=e_{1}, \quad\left[e_{3}, e_{1}\right]=-\left[e_{1}, e_{3}\right]=e_{2}
$$

Then $\mathbb{R}^{3}$ equipped with this bracket operation is a Lie algebra. In fact, as we will see later, this is the Lie algebra of (the matrix group) SO (3) and also of SU(2) in disguise.

Given two matrices $A, B \in \mathbb{k}^{n \times n}$, their commutator is

$$
[A, B]:=A B-B A
$$

$A$ and $B$ commute (i.e., $A B=B A$ ) if and only if $[A, B]=0$. The commutator $[\cdot, \cdot]$ is a product on $\mathbb{k}^{n \times n}$ satisfying conditions (LA1)-(LA3).
$\diamond$ Exercise 231 Verify the Jacobi identity for the commutator $[\cdot, \cdot]$.

The real vector space $\mathbb{k}^{n \times n}$ equipped with the commutator $[\cdot, \cdot]$ is a Lie algebra.

Note : The procedure to give $\mathbb{k}^{n \times n}$ a Lie algebra structure can be extended to any associative algebra. A Lie product (bracket) can be defined in any associative algebra by the comutator $[x, y]=x y-y x$, making it a Lie algebra. Here the skew-symmetry condition (axiom) is clearly satisfied, and one can check easily that in this case the Jacobi identity for the commutator follows from the associtivity law for the ordinary product.

There is another way in which Lie algebras arise in the study of algebras. A derivation $d$ of a nonassociative algebra $\mathcal{A}$ (i.e., a vector space endowed with a bilinear mapping $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ ) is a linear mapping $\mathcal{A} \rightarrow \mathcal{A}$ satisfying the formal analogue of the Leibniz rule for differentiating a product (for all $x, y \in \mathcal{A}$ )

$$
d(x y)=(d x) y+x(d y) .
$$

(The concept of a derivation is an abstraction of the idea of a first order differential operator.) The set of all derivations on $\mathcal{A}$ is clearly a vector subspace of the algebra End $(\mathcal{A})$ of all linear mappings $\mathcal{A} \rightarrow \mathcal{A}$. Although the product of derivations is in general not a derivation, the commutator $d_{1} d_{2}-d_{2} d_{1}$ of two derivations is again a derivation. Thus the set of all derivations of a nonassociative algebra is a Lie algebra, called the derivation algebra of the given nonassociative algebra.

Suppose that $\mathfrak{a}$ is a vector subspace of the Lie algebra $\mathbb{k}^{n \times n}$. Then $\mathfrak{a}$ is a Lie subalgebra of $\mathbb{k}^{n \times n}$ if it is closed under taking commutators of pairs of alements in $\mathfrak{a}$; that is,

$$
A, B \in \mathfrak{a} \Rightarrow[A, B] \in \mathfrak{a}
$$

Of course, $\mathbb{k}^{n \times n}$ is a Lie subalgebra of itself.
4.4.13 Theorem. If $G \leq G L(n, \mathbb{k})$ is a matrix group, then the tangent space $\mathfrak{g}=T_{I} G$ (at the identity) is a Lie subalgebra of $\mathbb{k}^{n \times n}$.

Proof : We will show that two curves $\alpha, \beta$ in $G$ with $\alpha(0)=\beta(0)=I$, there is such a curve $\gamma$ with $\dot{\gamma}(0)=[\dot{\alpha}(0), \dot{\beta}(0)]$.

Consider the mapping

$$
F:(s, t) \mapsto F(s, t):=\alpha(s) \beta(t) \alpha(s)^{-1}
$$

This is clearly (continuous and) differentiable with respect to each of the variables $s, t$. For each $s$ (in the domain of $\alpha$ ), $F(s, \cdot)$ is a curve in $G$ with $F(s, 0)=I$. Differentiating gives

$$
\left.\frac{d}{d t} F(s, t)\right|_{t=0}=\alpha(s) \dot{\beta}(0) \alpha(s)^{-1}
$$

and so

$$
\alpha(s) \dot{\beta}(0) \alpha(s)^{-1} \in \mathfrak{g} .
$$

Since $\mathfrak{g}$ is a closed subspace of $\mathbb{k}^{n \times n}$ (Any vector subspace is an intersection of hyperplanes), whenever this limit exists we also have

$$
\lim _{s \rightarrow 0} \frac{1}{s}\left(\alpha(s) \dot{\beta}(0) \alpha(s)^{-1}-\dot{\beta}(0)\right) \in \mathfrak{g}
$$

Exercise 232 Verify the following (matrix version of the) usual rule for differentiating an inverse :

$$
\frac{d}{d t}\left(\alpha(t)^{-1}\right)=-\alpha(t)^{-1} \dot{\alpha}(t) \alpha(t)^{-1}
$$

We have

$$
\begin{aligned}
\lim _{s \rightarrow 0} \frac{1}{s}\left(\alpha(s) \dot{\beta}(0) \alpha(s)^{-1}-\dot{\beta}(0)\right) & =\left.\frac{d}{d s} \alpha(s) \dot{\beta}(0) \alpha(s)^{-1}\right|_{s=0} \\
& =\dot{\alpha}(0) \dot{\beta}(0) \alpha(0)-\alpha(0) \dot{\beta}(0) \alpha(0)^{-1} \dot{\alpha}(0) \alpha(0)^{-1} \\
& =\dot{\alpha}(0) \dot{\beta}(0) \alpha(0)-\alpha(0) \dot{\beta}(0) \dot{\alpha}(0) \\
& =\dot{\alpha}(0) \dot{\beta}(0)-\dot{\beta}(0) \dot{\alpha}(0) \\
& =[\dot{\alpha}(0), \dot{\beta}(0)] .
\end{aligned}
$$

This shows that $[\dot{\alpha}(0), \dot{\beta}(0)] \in \mathfrak{g}$, hence it must be of the form $\dot{\gamma}(0)$ for some curve.

So for each matrix group $G$ there is a Lie algebra $\mathfrak{g}=T_{I} G$. We call $\mathfrak{g}$ the Lie algebra of $G$.

Note : The essential phenomenon behind Lie theory is that one may associate in a natural way to a matrix group $G$ its Lie algebra $\mathfrak{g}$. The Lie algebra is first of all a (real) vector space and secondly is endowed with a skew-symmetric bilinear product (called the Lie bracket or commutator). Amazingly, the group $G$ is almost completely determined by $\mathfrak{g}$ and its Lie bracket. Thus for many purposes one can replace $G$ with $\mathfrak{g}$. Since $G$ is a complicated nonlinear object and $\mathfrak{g}$ is just a vector space, it is usually vastly simpler to work with $\mathfrak{g}$. Otherwise intractable computations may become straightforward linear algebra. This is one source of the power of Lie theory.

## Homomorphisms of Lie algebras

A suitable type of homomorphism $G \rightarrow H$ between matrix groups gives rise to a linear mapping $\mathfrak{g} \rightarrow \mathfrak{h}$ respecting the Lie algebra structures.
4.4.14 Definition. Let $G \leq G L(n, \mathbb{k}), H \leq G L(m, \mathbb{k})$ be matrix groups and let $\varphi: G \rightarrow H$ be a continuous mapping. Then $\varphi$ is said to be differentiable if for every (differentiable) curve $\gamma:(a, b) \rightarrow G$, the composite mapping $\varphi \circ \gamma:(a, b) \rightarrow H$ is a (differentiable) curve with derivative

$$
(\varphi \circ \gamma)^{\cdot}(t)=\frac{d}{d t} \varphi(\gamma(t))
$$

and if whenever two (differentiable) curves $\alpha, \beta:(a, b) \rightarrow G$ both satisfy the conditions

$$
\alpha(0)=\beta(0) \quad \text { and } \quad \dot{\alpha}(0)=\dot{\beta}(0)
$$

then

$$
(\varphi \circ \alpha)^{\cdot}(0)=(\varphi \circ \beta)^{\cdot}(0)
$$

Such a mapping $\varphi$ is a differentiable homomorphism if it is also a group homomorphism. A continuous homomorphism of matrix groups that is also differentiable is called a Lie homomorphism.

Note : The "technical restriction" in the definition of a Lie homomorphism is in fact unnecessary.

If $\varphi: G \rightarrow H$ is the restriction of a differentiable mapping $\Phi: G \mathrm{~L}(n, \mathbb{k}) \rightarrow$ $\mathrm{GL}(m, \mathbb{k})$, then $\varphi$ is also a differentiable mapping.
4.4.15 Proposition. Let $G, H, K$ be matrix groups and $\varphi: G \rightarrow H, \psi$ : $H \rightarrow K$ be differentiable homomorphisms.
(a) For each $A \in G$ there is a linear mapping $d \varphi_{A}: T_{A} G \rightarrow T_{\varphi(A)} H$ given by

$$
d \varphi_{A}(\dot{\gamma}(0))=(\varphi \circ \gamma)^{\cdot}(0)
$$

(b) We have

$$
d \psi_{\varphi(A)} \circ d \varphi_{A}=d(\psi \circ \varphi)_{A}
$$

(c) For the identity mapping $i d_{G}: G \rightarrow G$ and $A \in G$,

$$
\operatorname{did}_{G}=i d_{T_{A} G}
$$

Proof: (a) The definition of $d \varphi_{A}$ makes sense since (by the definition of differentiability), given $X \in T_{A} G$, for any curve $\gamma$ with

$$
\gamma(0)=A \quad \text { and } \quad \dot{\gamma}(0)=X
$$

$(\varphi \circ \gamma)^{\cdot}(0)$ depends only on $X$ and not on $\gamma$.
$\diamond$ Exercise 233 Verify that the maping $d \varphi_{A}: T_{A} G \rightarrow T_{\varphi(A)} H$ is linear.
The identities (b) and (c) are straightforward to verify.
If $\varphi: G \rightarrow H$ is a differentiable homomorphism, then (since $\varphi(I)=I$ ) $d \varphi_{I}: T_{I} G \rightarrow T_{I} H$ is a linear mapping, called the derivative of $\varphi$ and usually denoted by

$$
d \varphi: \mathfrak{g} \rightarrow \mathfrak{h} .
$$

4.4.16 Definition. Let $\mathfrak{g}, \mathfrak{h}$ be Lie algebras. A linear mapping $\Phi: \mathfrak{g} \rightarrow \mathfrak{h}$ is a homomorphism of Lie algebras if (for $x, y \in \mathfrak{g}$ )

$$
\Phi([x, y])=[\Phi(x), \Phi(y)] .
$$

4.4.17 Theorem. Let $G, H$ be matrix groups and $\varphi: G \rightarrow H$ be a Lie homomorphism. Then the derivative $d \varphi: \mathfrak{g} \rightarrow \mathfrak{h}$ is a homomorphism of Lie algebras.

Proof : Following ideas and notation in the proof of Theorem 4.4.13, for curves $\alpha, \beta$ in $G$ with $\alpha(0)=\beta(0)=I$, we can use the composite mapping

$$
\varphi \circ F:(s, t) \mapsto \varphi(F(s, t))=\varphi(\alpha(s)) \varphi(\beta(t)) \varphi(\alpha(s))^{-1}
$$

to deduce

$$
d \varphi([\dot{\alpha}(0), \dot{\beta}(0)])=[d \varphi(\dot{\alpha}(0)), d \varphi(\dot{\beta}(0))] .
$$

4.4.18 Corollary. Let $G, H$ be matrix groups and $\varphi: G \rightarrow H$ be an isomorphism of matrix groups. Then the derivative $d \varphi: \mathfrak{g} \rightarrow \mathfrak{h}$ is an isomorphism of Lie algebras.

Proof : $\varphi^{-1} \circ \varphi$ is the identity, so

$$
d \varphi^{-1} \circ d \varphi: T_{I} G \rightarrow T_{I} G
$$

is the identity. Thus $d \varphi^{-1}$ is surjective and $d \varphi$ is injective.
Likewise, $\varphi \circ \varphi^{-1}$ is the identity, so $d \varphi \circ d \varphi^{-1}$ is the identity. Thus $d \varphi^{-1}$ is injective, and $d \varphi$ is surjective. The result now follows.

NOTE : Isomorphic matrix groups have isomorphic Lie algebras. The converse (i.e., matrix groups with isomorphic Lie algebras are isomorphic) is false. For example, the rotation group $\mathrm{SO}(2)$ and the diagonal group

$$
D_{1}=\left\{\left.\left[\begin{array}{cc}
1 & 0 \\
0 & e^{a}
\end{array}\right] \right\rvert\, a \in \mathbb{R}\right\} \leq \operatorname{GA}^{+}(1, \mathbb{R})
$$

have both Lie algebras isomorphic to $\mathbb{R}$ (the only Lie algebra structure on $\mathbb{R}$ ), but SO (2) is homeomorphic to a circle, while $D_{1}$ is homeomorphic to $\mathbb{R}$, so they are certainly not isomorphic.

However, the converse is - in a sense - almost true, so that the bracket operation on $\mathfrak{g}$ almost determine $G$ as a group. After the existence of the Lie algebra, this fact is the most remarkable in Lie theory. Its precise formulation is known as Lie's Third Theorem.

### 4.5 More Properties of the Exponential Mapping

The following formula can be considered as another definition of the matrix exponential.
4.5.1 Proposition. Let $A \in \mathbb{k}^{n \times n}$. Then

$$
\exp (A)=\lim _{r \rightarrow \infty}\left(I+\frac{1}{r} A\right)^{r}
$$

Proof: Consider the difference

$$
\exp (A)-\left(I+\frac{1}{r} A\right)^{r}=\sum_{k=0}^{\infty}\left(\frac{1}{k!}-\frac{1}{r^{k}}\binom{r}{k}\right) A^{k} .
$$

This matrix series converges since the series for the matrix exponential $\exp (A)$ converges and $\left(I+\frac{1}{r} A\right)^{r}$ is a polynomial. The coefficients in the rhs are nonnegative since

$$
\frac{1}{k!} \geq \frac{r(r-1) \cdots(r-k+1)}{r \cdot r \cdots r} \frac{1}{k!} .
$$

Therefore, setting $\|A\|=a$, we get

$$
\left\|\exp (A)-\left(I+\frac{1}{r} A^{r}\right)^{r}\right\| \leq \sum_{k=0}^{\infty}\left(\frac{1}{k!}-\frac{1}{r^{k}}\binom{r}{k}\right) a^{k}=e^{a}-\left(1+\frac{a}{r}\right)^{r}
$$

where the expression on the right approaches zero (as $r \rightarrow \infty$ ). The result now follows.
4.5.2 Proposition. Let $A \in \mathbb{k}^{n \times n}$ and $\epsilon \in \mathbb{R}$. Then

$$
\operatorname{det}(I+\epsilon A)=1+\epsilon \operatorname{tr} A+O\left(\epsilon^{2}\right) \quad(\text { as } \epsilon \rightarrow 0) .
$$

Proof: The determinant of $I+\epsilon A$ equals the product of the eigenvalues of the matrix. But the eigenvalues of $I+\epsilon A$ (with due regard for multiplicity) equal $1+\epsilon \lambda_{i}$, where the $\lambda_{i}$ are the eigenvalues of $A$. It follows that

$$
\begin{aligned}
\operatorname{det}(I+\epsilon A) & =\left(1+\epsilon \lambda_{1}\right)\left(1+\epsilon \lambda_{2}\right) \cdots\left(1+\epsilon \lambda_{n}\right) \\
& =1+\epsilon\left(\lambda_{1}+\lambda_{2}+\cdots+\lambda_{n}\right)+O\left(\epsilon^{2}\right) \\
& =1+\epsilon \operatorname{tr} A+O\left(\epsilon^{2}\right) .
\end{aligned}
$$

Note : Whenever we have a mapping $Z$ from some (open) interval $(a, b), a<$ $0<b$ into a finite-dimensional normed vector space (e.g. $\mathbb{k}^{n \times n}$ ), then $Z$ will often be denoted by $O\left(t^{k}\right)$ if $t \mapsto \frac{1}{t^{k}} Z(t)$ is bounded in an (open) neighborhood of the origin 0 (i.e. there are constants $C_{1}$ and $C_{2}$ such that

$$
\left.\|Z(t)\| \leq C_{1}\left|t^{k}\right| \quad \text { for }|t|<C_{2} .\right)
$$

Thus $O\left(t^{k}\right)$ may denote different mappings at different times. The big- $O$ notation was first introduced in 1892 by Paul G.H. Bachmann (1837-1920) in a book on number theory, and is currently used in several areas of mathematics and computer science (including mathematical analysis and the theory of algorithms).

The following result is useful.
4.5.3 Lemma. Let $\alpha:(a, b) \rightarrow \mathbb{k}^{n \times n}$ be a curve. Then

$$
\left.\frac{d}{d t} \operatorname{det} \alpha(t)\right|_{t=0}=\operatorname{tr} \dot{\alpha}(0)
$$

Proof : The operation $\partial:=\left.\frac{d}{d t}\right|_{t=0}$ has the derivation property

$$
\partial\left(\gamma_{1} \gamma_{2}\right)=\left(\partial \gamma_{1}\right) \gamma_{2}(0)+\gamma_{1}(0) \partial \gamma_{2}
$$

Put $\alpha(t)=\left[a_{i j}\right]$ and notice that (when $\left.t=0\right) a_{i j}=\delta_{i j}$. Write $C_{i j}$ for the cofactor matrix obtained from $\alpha(t)$ by deleting the $i^{\text {th }}$ row and the $j^{\text {th }}$ column. By expanding along the $n^{\text {th }}$ row we get

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

For $t=0($ since $\alpha(0)=I)$ we have

$$
\operatorname{det} C_{n j}=\delta_{n j}
$$

Then

$$
\begin{aligned}
\partial \operatorname{det} \alpha(t) & =\sum_{j=1}^{n}(-1)^{n+j}\left(\left(\partial a_{n j}\right) \operatorname{det} C_{n j}+a_{n j}\left(\partial \operatorname{det} C_{n j}\right)\right) \\
& =\sum_{j=1}^{n}(-1)^{n+j}\left(\left(\partial a_{n j}\right) \operatorname{det} C_{n j}\right)+\left(\partial \operatorname{det} C_{n n}\right) \\
& =\partial a_{n n}+\partial \operatorname{det} C_{n n}
\end{aligned}
$$

We can repeat this calculation with the $(n-1) \times(n-1)$ matrix $C_{n n}$ and so on. This gives

$$
\begin{aligned}
\partial \operatorname{det} \alpha(t) & =\partial a_{n n}+\partial a_{n-1, n-1}+\partial \operatorname{det} C_{n-1, n-1} \\
& \vdots \\
& =\partial a_{n n}+\partial a_{n-1, n-1}+\cdots+\partial a_{11} \\
& =\operatorname{tr} \dot{\alpha}(0) .
\end{aligned}
$$

We can now prove a remarkable (and very useful) result, known as Liouville's Formula. Three different proofs will be given.
4.5.4 Theorem. (Liouville's Formula) For $A \in \mathbb{k}^{n \times n}$ we have

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

First solution (using the second definition of the exponential): We have

$$
\operatorname{det} \exp (A)=\operatorname{det} \lim _{r \rightarrow \infty}\left(I+\frac{1}{r} A\right)^{r}=\lim _{r \rightarrow \infty} \operatorname{det}\left(I+\frac{1}{r} A\right)^{r}
$$

since the determinant function det: $\mathbb{k}^{n \times n} \rightarrow \mathbb{k}$ is continuous. Moreover, by Proposition 4.5.2,

$$
\operatorname{det}\left(I+\frac{1}{r} A\right)^{r}=\left[\operatorname{det}\left(I+\frac{1}{r} A\right)\right]^{r}=\left[1+\frac{1}{r} \operatorname{tr} A+O\left(\frac{1}{r^{2}}\right)\right]^{r}(\text { as } r \rightarrow \infty) .
$$

It only remains to note that (for any $a \in \mathbb{k}$ )

$$
\lim _{r \rightarrow \infty}\left[1+\frac{a}{r}+O\left(\frac{1}{r^{2}}\right)\right]^{r}=e^{a} .
$$

In particular, for $a=\operatorname{tr} A$, we get the desired result.
Second solution (using differential equations) : Consider the curve

$$
\gamma: \mathbb{R} \rightarrow \mathrm{GL}(1, \mathbb{k})=\mathbb{k}^{\times}, \quad t \mapsto \operatorname{det} \exp (t A) .
$$

Then (by Lemma 4.5.3 applied to the curve $\gamma$ )

$$
\begin{aligned}
\dot{\gamma}(t) & =\lim _{h \rightarrow 0} \frac{1}{h}[\operatorname{det} \exp ((t+h) A)-\operatorname{det} \exp (t A)] \\
& =\operatorname{det} \exp (t A) \lim _{h \rightarrow 0} \frac{1}{h}[\operatorname{det} \exp (h A)-1] \\
& =\operatorname{det} \exp (t A) \operatorname{tr} A \\
& =\gamma(t) \operatorname{tr} A .
\end{aligned}
$$

So $\gamma$ satisfies the same differential equation and initial condition as the curve $t \mapsto e^{t \operatorname{tr} A}$. By the uniqueness of the solution (see Exercise 219), it follows that

$$
\gamma(t)=\operatorname{det} \exp (t A)=e^{t \operatorname{tr} A} .
$$

In particular, for $t=1$, we get the desired result.
Third solution (using Jordan canonical form) : If $B \in G L(n, \mathbb{k})$, then (see Exercise 221)

$$
\begin{aligned}
\operatorname{det} \exp \left(B A B^{-1}\right) & =\operatorname{det}\left(B \exp (A) B^{-1}\right) \\
& =\operatorname{det} B \cdot \operatorname{det} \exp (A) \cdot \operatorname{det} B^{-1} \\
& =\operatorname{det} \exp (A)
\end{aligned}
$$

and

$$
e^{\operatorname{tr}\left(B A B^{-1}\right)}=e^{\operatorname{tr} A}
$$

So it suffices to prove the identity for $B A B^{-1}$ for a suitably chosen invertible matrix $B$. Using for example the theory of Jordan canonical forms, there is a suitable choice of such a $B$ for which

$$
B A B^{-1}=D+N
$$

with $D$ diagonal and $N$ strictly upper triangular (i.e., $N_{i j}=0$ for $i \geq j$ ). Then $N$ is nilpotent (i.e., $N^{k}=O$ for some $k \geq 1$ ). We have

$$
\begin{aligned}
\exp \left(B A B^{-1}\right) & =\sum_{k=0}^{\infty} \frac{1}{k!}(D+N)^{k} \\
& =\sum_{k=0}^{\infty} \frac{1}{k!} D^{k}+\sum_{k=0}^{\infty} \frac{1}{(k+1)!}\left((D+N)^{k+1}-D^{k+1}\right) \\
& =\exp (D)+\sum_{k=0}^{\infty} \frac{1}{(k+1)!} N\left(D^{k}+D^{k-1} N+\cdots+N^{k}\right)
\end{aligned}
$$

The matrix

$$
N\left(D^{k}+D^{k-1} N+\cdots+N^{k}\right)
$$

is strictly upper triangular, and so

$$
\exp \left(B A B^{-1}\right)=\exp (D)+N^{\prime}
$$

where $N^{\prime}$ is strictly upper triangular. Now, if $D=\operatorname{diag}\left(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{n}\right)$, we
have

$$
\begin{aligned}
\operatorname{det} \exp (A) & =\operatorname{det} \exp \left(B A B^{-1}\right) \\
& =\operatorname{det} \exp (D) \\
& =\operatorname{det} \operatorname{diag}\left(e^{\lambda_{1}}, e^{\lambda_{2}}, \ldots, e^{\lambda_{n}}\right) \\
& =e^{\lambda_{1}} e^{\lambda_{2}} \cdots e^{\lambda_{n}} \\
& =e^{\lambda_{1}+\lambda_{2}+\cdots+\lambda_{n}} \\
& =e^{\operatorname{tr} D} \\
& =e^{\operatorname{tr}\left(B A B^{-1}\right)} \\
& =e^{\operatorname{tr} A}
\end{aligned}
$$

The exponential mapping

$$
\exp : \mathbb{k}^{n \times n} \rightarrow \mathrm{GL}(n, \mathbb{k})
$$

is a basic link between the linear structure on $\mathbb{k}^{n \times n}$ and the multiplicative structure on $\mathrm{GL}(n, \mathbb{k})$. Let $G$ be a matrix subgroup of $\mathrm{GL}(n, \mathbb{k})$. Applying Proposition 4.3.3, we may choose $\rho \in \mathbb{R}$ so that $0<\rho \leq \frac{1}{2}$ and if $A, B \in$ $\mathcal{B}_{\mathbb{k}^{n \times n}}(O, \rho)$, then $\exp (A) \exp (B) \in \exp \left(\mathcal{B}_{\mathbb{k}^{n \times n}}\left(O, \frac{1}{2}\right)\right)$. Since exp is one-toone on $\mathcal{B}_{\mathbb{k}^{n \times n}}(O, \rho)$, there is a unique matrix $C \in \mathbb{k}^{n \times n}$ for which

$$
\exp (A) \exp (B)=\exp (C) .
$$

NOTE : There is a beautiful formula, the Baker-Campbell-Hausdorff formula which expresses $C$ as a universal power series in $A$ and $B$. To develop this completely would take too long. Specifically, (one form of) the B-C-H formula says that if $X$ and $Y$ are sufficiently small, then

$$
\begin{gathered}
\exp (X) \exp (Y)=\exp (Z) \text { with } \\
Z=X+Y+\frac{1}{2}[X, Y]+\frac{1}{12}[X,[X, Y]]-\frac{1}{12}[Y,[X, Y]]+\cdots
\end{gathered}
$$

It is not supposed to be evident at the moment what "..." refers to. The only important point is that all the terms (in the expansion of $Z$ ) are given in terms of $X$
and $Y$, Lie brackets of $X$ and $Y$, Lie brackets of Lie brackets involving $X$ and $Y$, etc. Then it follows that the mapping $\phi: G \rightarrow G L(n, \mathbb{R})$ "defined" by the relation

$$
\phi(\exp (X))=\exp (\phi(X))
$$

is such that on elements of the form $\exp (X)$, with $X$ sufficiently small, is a group homomorphism. Hence the B-C-H formula shows that all the information about the group product, a least near the identity, is "encoded" in the Lie algebra.

An interesting special case is the following : If $X, Y \in \mathbb{C}^{n \times n}$ and $X, Y$ commute with their commutator (i.e., $[X,[X, Y]]=[Y,[X, Y]$ ), then

$$
\exp (X) \exp (Y)=\exp \left(X+Y+\frac{1}{2}[X, Y]\right)
$$

Exercise 234 Show by direct computation that for

$$
X, Y \in \mathfrak{h e i s}=\left\{\left.\left[\begin{array}{ccc}
0 & a & b \\
0 & 0 & c \\
0 & 0 & 0
\end{array}\right] \right\rvert\, a, b, c \in \mathbb{R}\right\}
$$

(the Lie algebra of the Heisenberg group Heis)

$$
\exp (X) \exp (Y)=\exp \left(X+Y+\frac{1}{2}[X, Y]\right)
$$

We set

$$
R=C-A-B \in \mathbb{k}^{n \times n} .
$$

For $X \in \mathbb{k}^{n \times n}$, we have

$$
\exp (X)=I+X+R_{1}(X)
$$

where the remainder term $R_{1}(X)$ is given by

$$
R_{1}(X)=\sum_{k=2}^{\infty} \frac{1}{k!} X^{k}
$$

Hence

$$
\left\|R_{1}(X)\right\| \leq\|X\|^{2} \sum_{k=2}^{\infty} \frac{1}{k!}\|X\|^{k-2}
$$

and therefore if $\|X\|<1$, then

$$
\left\|R_{1}(X)\right\| \leq\|X\|^{2} \sum_{k=2}^{\infty} \frac{1}{k!}=\|X\|^{2}(e-2)<\|X\|^{2} .
$$

Now for $X=C \in \mathcal{B}_{\mathbb{k}^{n \times n}}\left(O, \frac{1}{2}\right)$, we have

$$
\exp (C)=I+C+R_{1}(C)
$$

with

$$
\left\|R_{1}(C)\right\|<\|C\|^{2} .
$$

Similar considerations lead to

$$
\exp (C)=\exp (A) \exp (B)=I+A+B+R_{1}(A, B),
$$

where

$$
R_{1}(A, B)=\sum_{k=2}^{\infty} \frac{1}{k!}\left(\sum_{r=0}^{k}\binom{k}{r} A^{r} B^{k-r}\right) .
$$

This gives

$$
\begin{aligned}
\left\|R_{1}(A, B)\right\| & \leq \sum_{k=2}^{\infty} \frac{1}{k!}\left(\sum_{r=0}^{k}\binom{k}{r}\|A\|^{r}\|B\|^{k-r}\right) \\
& =\sum_{k=0}^{\infty} \frac{1}{k!}(\|A\|+\|B\|)^{k} \\
& =(\|A\|+\|B\|)^{2} \sum_{k=2}^{\infty} \frac{1}{k!}(\|A\|+\|B\|)^{k-2} \\
& \leq(\|A\|+\|B\|)^{2}
\end{aligned}
$$

since $\|A\|+\|B\|<1$.
Combining the two ways of writing $\exp (C)$ from above, we have

$$
C=A+B+R_{1}(C)-R_{1}(A, B)
$$

and so

$$
\begin{aligned}
\|C\| & \leq\|A\|+\|B\|+\left\|R_{1}(A, B)\right\|+\left\|R_{1}(C)\right\| \\
& <\|A\|+\|B\|+(\|A\|+\|B\|)^{2}+\|C\|^{2} \\
& \leq 2(\|A\|+\|B\|)+\frac{1}{2}\|C\|
\end{aligned}
$$

since $\|A\|,\|B\|,\|C\| \leq \frac{1}{2}$. Finally this gives

$$
\|C\| \leq 4(\|A\|+\|B\|) .
$$

We also have

$$
\begin{aligned}
\|R\|=\|C-A-B\| & \leq\left\|R_{1}(A, B)\right\|+\left\|R_{1}(C)\right\| \\
& \leq(\|A\|+\|B\|)^{2}+(4(\|A\|+\|B\|))^{2} \\
& =17(\|A\|+\|B\|)^{2} .
\end{aligned}
$$

We have proved the following result.
4.5.5 Proposition. Let $A, B, C \in \mathcal{B}_{\mathbb{k}^{n \times n}}\left(O, \frac{1}{2}\right)$ such that $\exp (A) \exp (B)=$ $\exp (C)$. Then $C=A+B+R$, where the remainder term $R$ satisfies

$$
\|R\| \leq 17(\|A\|+\|B\|)^{2}
$$

We can refine this estimate (to second order). We only point out the essential steps (details will be omitted). Set

$$
S=C-A-B-\frac{1}{2}[A, B] \in \mathbb{K}^{n \times n}
$$

and write

$$
\exp (C)=I+C+\frac{1}{2} C^{2}+R_{2}(C)
$$

with

$$
\left\|R_{2}(C)\right\| \leq \frac{1}{3}\|C\|^{3} .
$$

Then

$$
\begin{aligned}
\exp (C) & =I+A+B+\frac{1}{2}[A, B]+S+\frac{1}{2} C^{2}+R_{2}(C) \\
& =I+A+B+\frac{1}{2}\left(A^{2}+2 A B+B^{2}\right)+T,
\end{aligned}
$$

where

$$
T=S+\frac{1}{2}\left(C^{2}-(A+B)^{2}\right)+R_{2}(C)
$$

Also

$$
\exp (A) \exp (B)=I+A+B+\frac{1}{2}\left(A^{2}+2 A B+B^{2}\right)+R_{2}(A, B)
$$

with

$$
\left\|R_{2}(A, B)\right\| \leq \frac{1}{3}(\|A\|+\|B\|)^{3} .
$$

We see that

$$
S=R_{2}(A, B)+\frac{1}{2}\left((A+B)^{2}-C^{2}\right)-R_{2}(C)
$$

and by taking norms we get

$$
\begin{aligned}
\|S\| & \leq\left\|R_{2}(A, B)\right\|+\frac{1}{2}\|(A+B)(A+B-C)+(A+B-C) C\|+\left\|R_{2}(C)\right\| \\
& \leq \frac{1}{3}(\|A\|+\|B\|)^{3}+\frac{1}{2}(\|A\|+\|B\|+\|C\|)\|A+B-C\|+\frac{1}{3}\|C\|^{3} \\
& \leq 65(\|A\|+\|B\|)^{3} .
\end{aligned}
$$

The following estimation holds.
4.5.6 Proposition. Let $A, B, C \in \mathcal{B}_{\mathbb{k}^{n \times n}}\left(O, \frac{1}{2}\right)$ such that $\exp (A) \exp (B)=$ $\exp (C)$. Then $C=A+B+\frac{1}{2}[A, B]+S$, where the remainder term $S$ satisfies

$$
\|S\| \leq 65(\|A\|+\|B\|)^{3}
$$

We will derive two main consequences of Proposition 4.5.5 and PropoSition 4.5.6. These relate group operations in $\mathrm{GL}(n, \mathbb{k})$ to the linear operations in $\mathbb{k}^{n \times n}$ and are crucial ingredients in the proof that every matrix group is a Lie group.
4.5.7 Theorem. (Lie-Trotter Product Formula) For $U, V \in \mathbb{k}^{n \times n}$ we have

$$
\exp (U+V)=\lim _{r \rightarrow \infty}\left(\exp \left(\frac{1}{r} U\right) \exp \left(\frac{1}{r} V\right)\right)^{r}
$$

(This formula relates addition in $\mathbb{k}^{n \times n}$ to multiplication in $\mathrm{GL}(n, \mathbb{k})$.)
Proof : For large $r$ we may take $A=\frac{1}{r} U$ and $B=\frac{1}{r} V$ and apply PropoSItion 4.5.5 to give

$$
\exp \left(\frac{1}{r} U\right) \exp \left(\frac{1}{r} V\right)=\exp \left(C_{r}\right)
$$

with

$$
\left\|C_{r}-\frac{1}{r}(U+V)\right\| \leq \frac{17(\|U\|+\|V\|)^{2}}{r^{2}}
$$

As $r \rightarrow \infty$,

$$
\left\|r C_{r}-(U+V)\right\| \leq \frac{17(\|U\|+\|V\|)^{2}}{r} \rightarrow 0
$$

and hence

$$
r C_{r} \rightarrow U+V
$$

Since $\exp \left(r C_{r}\right)=\exp \left(C_{r}\right)^{r}$, the Lie-Trotter product formula follows by continuity of the exponential mapping.
4.5.8 Theorem. (Commutator Formula) For $U, V \in \mathbb{k}^{n \times n}$ we have

$$
\exp ([U, V])=\lim _{r \rightarrow \infty}\left(\exp \left(\frac{1}{r} U\right) \exp \left(\frac{1}{r} V\right) \exp \left(-\frac{1}{r} U\right) \exp \left(-\frac{1}{r} V\right)\right)^{r^{2}} .
$$

(This formula relates the Lie bracket - or commutator - in $\mathbb{k}^{n \times n}$ to the group commutator in $G L(n, \mathbb{k})$.)

Proof : For large $r$ (as in the proof of Theorem 4.5.7) we have

$$
\exp \left(\frac{1}{r} U\right) \exp \left(\frac{1}{r} V\right)=\exp \left(C_{r}\right)
$$

with (as $r \rightarrow \infty$ )

$$
r C_{r} \rightarrow U+V
$$

We also have

$$
C_{r}=\frac{1}{r}(U+V)+\frac{1}{2 r^{2}}[U, V]+S_{r}
$$

where

$$
\left\|S_{r}\right\| \leq 65 \frac{(\|U\|+\|V\|)^{3}}{r^{3}}
$$

Similarly (replacing $U, V$ with $-U,-V$ ) we obtain :

$$
\exp \left(-\frac{1}{r} U\right) \exp \left(-\frac{1}{r} V\right)=\exp \left(C_{r}^{\prime}\right)
$$

where

$$
C_{r}^{\prime}=-\frac{1}{r}(U+V)+\frac{1}{2 r^{2}}[U, V]+S_{r}^{\prime}
$$

and

$$
\left\|S_{r}^{\prime}\right\| \leq 65 \frac{(\|U\|+\|V\|)^{3}}{r^{3}}
$$

Combining these we get

$$
\begin{aligned}
\exp \left(\frac{1}{r} U\right) \exp \left(\frac{1}{r} V\right) \exp \left(-\frac{1}{r} U\right) \exp \left(-\frac{1}{r} V\right) & =\exp \left(C_{r}\right) \exp \left(C_{r}^{\prime}\right) \\
& =\exp \left(E_{r}\right)
\end{aligned}
$$

where

$$
\begin{aligned}
E_{r} & =C_{r}+C_{r}^{\prime}+\frac{1}{2}\left[C_{r}, C_{r}^{\prime}\right]+T_{r} \\
& =\frac{1}{r^{2}}[U, V]+\frac{1}{2}\left[C_{r}, C_{r}^{\prime}\right]+S_{r}+S_{r}^{\prime}+T_{r}
\end{aligned}
$$

$\diamond$ Exercise 235 Verify that

$$
\begin{aligned}
{\left[C_{r}, C_{r}^{\prime}\right]=} & \frac{1}{r^{3}}[U+V,[U, V]]+\frac{1}{r}\left[U+V, S_{r}+S_{r}^{\prime}\right] \\
& +\frac{1}{2 r^{2}}\left[[U, V], S_{r}^{\prime}-S_{r}\right]+\left[S_{r}, S_{r}^{\prime}\right]
\end{aligned}
$$

All four of these terms have norm bounded by an expression of the form $\frac{\text { constant }}{r^{3}}$ so the same is true of $\left[C_{r}, C_{r}^{\prime}\right]$. Also $S_{r}, S_{r}^{\prime}, T_{r}$ have similarly bounded norms. Setting

$$
Q_{r}:=r^{2} E_{r}-[U, V]
$$

we obtain (as $r \rightarrow \infty$ )

$$
\left\|Q_{r}\right\|=r^{2}\left\|E_{r}-\frac{1}{r^{2}}[U, V]\right\| \leq \frac{\text { constant }}{r} \rightarrow 0
$$

and hence

$$
\exp \left(E_{r}\right)^{r^{2}}=\exp \left([U, V]+Q_{r}\right) \rightarrow \exp ([U, V])
$$

The commutator formula now follows using continuity of the exponential mapping.

Note : If $g, h$ are elements of a group, then the expression $g h g^{-1} h^{-1}$ is called the group commutator of $g$ and $h$.

There is one further concept involving the exponential mapping that is basic in Lie theory. It involves conjugation, which is generally referred to as the "adjoint action". For $g \in G \mathrm{~L}(n, \mathbb{k})$ and $A \in \mathbb{k}^{n \times n}$, we can form the conjugate

$$
\operatorname{Ad}_{g}(A):=g A g^{-1}
$$

$\diamond$ Exercise 236 Let $A, B \in \mathbb{k}^{n \times n}$ and $g, h \in \mathrm{GL}(n, \mathbb{k})$. Show that (for $\lambda, \mu \in \mathbb{k}$ )
(a) $\operatorname{Ad}_{g}(\lambda A+\mu B)=\lambda \operatorname{Ad}_{g}(A)+\mu \operatorname{Ad}_{g}(B)$.
(b) $\operatorname{Ad}_{g}([A, B])=\left[\operatorname{Ad}_{g}(A), \operatorname{Ad}_{g}(B)\right]$.
(c) $\operatorname{Ad}_{g h}(A)=\operatorname{Ad}_{g}\left(\operatorname{Ad}_{h}(A)\right)$.

In particular, $\operatorname{Ad}_{g}^{-1}=\operatorname{Ad}_{g^{-1}}$.
Formulas (a) an (b) say that $\operatorname{Ad}_{g}$ is an automorphism of the Lie algebra $\mathbb{k}^{n \times n}$, and formula (c) says the mapping

$$
\operatorname{Ad}: \mathrm{GL}(n, \mathbb{k}) \rightarrow \operatorname{Aut}\left(\mathbb{k}^{n \times n}\right), \quad g \mapsto \operatorname{Ad}_{g}
$$

is a group homomorphism. The mapping Ad is called the adjoint representation of (the matrix group) $\mathrm{GL}(n, \mathbb{k})$.

Formula $(c)$ implies in particular that if $t \mapsto \exp (t A)$ is a one-parameter subgroup of $\mathrm{GL}(n, \mathbb{k})$, then $\operatorname{Ad}_{\exp (t A)}$ is a one-parameter group (of linear transformations) in $\mathbb{k}^{n \times n}$. Observe that we can identify Aut ( $\mathbb{k}^{n \times n}$ ) with $\mathrm{GL}\left(n^{2}, \mathbb{k}\right)$ (and thus view Aut $\left(\mathbb{k}^{n \times n}\right.$ ) as a matrix group). Then (by THEOREM 4.4.3)

$$
\operatorname{Ad}_{\exp (t A)}=\exp (t \mathcal{A})
$$

for some $\mathcal{A} \in \mathbb{k}^{n^{2} \times n^{2}}=$ End $\left(\mathbb{k}^{n \times n}\right)$. Since

$$
\begin{aligned}
\mathcal{A}(B) & =\left.\frac{d}{d t} \operatorname{Ad}_{\exp (t A)}(B)\right|_{t=0} \\
& =\left.\frac{d}{d t} \exp (t A) B \exp (-t A)\right|_{t=0} \\
& =[A, B]
\end{aligned}
$$

by setting (for $A, B \in \mathbb{K}^{n \times n}$ )

$$
\operatorname{ad} A(B):=[A, B]
$$

we have the following formula

$$
\operatorname{Ad}_{\exp (t A)}=\exp (t \operatorname{ad} A)
$$

Explicitly, the formula says that

$$
\exp (t A) B \exp (-t A)=\sum_{k=0}^{\infty} \frac{t^{k}}{k!}(\operatorname{ad} A)^{k} B
$$

(Here $(\operatorname{ad} A)^{0}=A$ and $(\operatorname{ad} A)^{k}=\operatorname{ad}(\operatorname{ad} A)^{k-1}$ for $k \geq 1$.)
Note: The mapping

$$
\operatorname{ad}: \mathbb{k}^{n \times n} \rightarrow \operatorname{End}\left(\mathbb{k}^{n \times n}\right), \quad X \mapsto \operatorname{ad} X
$$

is called the adjoint representation of (the Lie algebra) $\mathbb{k}^{n \times n}$. From the Jacobi identity for Lie algebras, we have

$$
\operatorname{ad} X([Y, Z])=[\operatorname{ad} X(Y), Z]+[Y, \operatorname{ad} X(Z)] .
$$

That is, ad $X$ is a derivation of the Lie algebra $\mathbb{k}^{n \times n}$. The formula above gives the relation between the automorphism $\operatorname{Ad}_{\exp (t X)}$ of the Lie algebra $\mathbb{k}^{n \times n}$ and the derivation $\operatorname{ad} X$ of $\mathbb{k}^{n \times n}$. One also has

$$
\exp \left(t \operatorname{Ad}_{g}(X)\right)=g \exp (t X) g^{-1}
$$

Using this formula, we can see that $[X, Y]=0$ if and only if $\exp (t X)$ and $\exp (s Y)$ commute for arbitrary $s, t \in \mathbb{R}$.
$\diamond$ Exercise 237 Let $A, B \in \mathbb{k}^{n \times n}$.
(a) Verify that

$$
\operatorname{ad}[A, B]=\operatorname{ad} A \operatorname{ad} B-\operatorname{ad} B \operatorname{ad} A=[\operatorname{ad} A, \operatorname{ad} B] .
$$

(This means that ad : $\mathbb{k}^{n \times n} \rightarrow \operatorname{End}\left(\mathbb{k}^{n \times n}\right)$ is a Lie algebra homomorphism.)
(b) Show by induction that

$$
(\operatorname{ad} A)^{n}(B)=\sum_{k=0}^{n}\binom{n}{k} A^{k} B(-A)^{n-k} .
$$

(c) Show by direct computation that

$$
\exp (\operatorname{ad} A)(B)=\operatorname{Ad}_{\exp (A)}(B)=\exp (A) B \exp (-A)
$$

### 4.6 Examples of Lie Algebras of Matrix Groups

## The Lie algebras of $G L(n, \mathbb{R})$ and $G L(n, \mathbb{C})$

Let us start with the real general linear group $G L(n, \mathbb{R}) \subseteq \mathbb{R}^{n \times n}$. We have shown (see EXAMPLE 4.4.8) that $T_{I} \mathrm{GL}(n, \mathbb{R})=\mathbb{R}^{n \times n}$. Hence the Lie algebra $\mathfrak{g l}(n, \mathbb{R})$ of $\mathrm{GL}(n, \mathbb{R})$ consists of all $n \times n$ matrices (with real entries), with the commutator as the Lie bracket. Thus

$$
\mathfrak{g l}(n, \mathbb{R})=\mathbb{R}^{n \times n}
$$

It follows that

$$
\operatorname{dim} \mathrm{GL}(n, \mathbb{R})=\operatorname{dim} \mathfrak{g l}(n, \mathbb{R})=n^{2}
$$

Similarly, the Lie algebra of the complex general linear group $\mathrm{GL}(n, \mathbb{C})$ is

$$
\mathfrak{g l}(n, \mathbb{C})=\mathbb{C}^{n \times n}
$$

and

$$
\operatorname{dim} \mathrm{GL}(n, \mathbb{C})=\operatorname{dim}_{\mathbb{R}} \mathfrak{g l}(n, \mathbb{C})=2 n^{2}
$$

## The Lie algebras of $\operatorname{SL}(n, \mathbb{R})$ and $\operatorname{SL}(n, \mathbb{C})$

For $S L(n, \mathbb{R}) \leq G L(n, \mathbb{R})$, suppose that

$$
\alpha:(a, b) \rightarrow \operatorname{SL}(n, \mathbb{R})
$$

is a curve in $\operatorname{SL}(n, \mathbb{R})$ with $\alpha(0)=I$. For $t \in(a, b)$ we have $\operatorname{det} \alpha(t)=1$ and so

$$
\frac{d}{d t} \operatorname{det} \alpha(t)=0
$$

Using Lemma 4.5.3, it follows that

$$
\operatorname{tr} \dot{\alpha}(0)=0
$$

and thus

$$
T_{I} S L(n, \mathbb{R}) \subseteq \operatorname{ker} \operatorname{tr}:=\left\{A \in \mathbb{R}^{n \times n} \mid \operatorname{tr} A=0\right\}
$$

If $A \in \operatorname{ker} \operatorname{tr} \subseteq \mathbb{R}^{n \times n}$, the curve

$$
\alpha:(a, b) \rightarrow \mathbb{R}^{n \times n}, \quad t \mapsto \exp (t A)
$$

satisfies (the boundary conditions)

$$
\alpha(0)=I \quad \text { and } \quad \dot{\alpha}(0)=A .
$$

Moreover, using Liouville's Formula, we get

$$
\operatorname{det} \alpha(t)=\operatorname{det} \exp (t A)=e^{t \operatorname{tr} A}=1
$$

Hence the Lie algebra $\mathfrak{s l}(n, \mathbb{R})$ of $\operatorname{SL}(n, \mathbb{R})$ consists of all $n \times n$ matrices (with real entries) having trace zero, with the commutator as the Lie bracket. Thus

$$
\mathfrak{s l}(n, \mathbb{R})=T_{I} \mathrm{SL}(n, \mathbb{R})=\{A \in \mathfrak{g l}(n, \mathbb{R}) \mid \operatorname{tr} A=0\}
$$

Since $\operatorname{tr} A=0$ imposes one condition on $A$, it follows that

$$
\operatorname{dim} \operatorname{SL}(n, \mathbb{R})=\operatorname{dim}_{\mathbb{R}} \mathfrak{s l}(n, \mathbb{R})=n^{2}-1
$$

Similarly, the Lie algebra of the complex special linear group $\operatorname{SL}(n, \mathbb{C})$ is

$$
\mathfrak{s l}(n, \mathbb{C})=T_{I} \mathrm{SL}(n, \mathbb{C})=\{A \in \mathfrak{g l}(n, \mathbb{C}) \mid \operatorname{tr} A=0\}
$$

and

$$
\operatorname{dim} \operatorname{SL}(n, \mathbb{C})=\operatorname{dim}_{\mathbb{R}} \mathfrak{s l}(n, \mathbb{C})=2\left(n^{2}-1\right) .
$$

The Lie algebras of $\mathrm{O}(n)$ and $\mathrm{SO}(n)$
First, consider the orthogonal group $\mathrm{O}(n)$; that is,

$$
\mathrm{O}(n)=\left\{A \in \mathrm{GL}(n, \mathbb{R}) \mid A^{\top} A=I\right\} \leq \mathrm{GL}(n, \mathbb{R}) .
$$

Given a curve $\alpha:(a, b) \rightarrow \mathrm{O}(n)$ with $\alpha(0)=I$, we have

$$
\frac{d}{d t} \alpha(t)^{T} \alpha(t)=0
$$

and so

$$
\dot{\alpha}(t)^{T} \alpha(t)+\alpha(t)^{T} \dot{\alpha}(t)=0
$$

which implies

$$
\dot{\alpha}(0)^{T}+\dot{\alpha}(0)=0 .
$$

Thus we must have $\dot{\alpha}(0) \in \mathbb{R}^{n \times n}$ is skew-symmetric. So

$$
T_{I} \mathrm{O}(n) \subseteq \operatorname{Sk}-\operatorname{sym}(n)=\left\{A \in \mathbb{R}^{n \times n} \mid A^{\top}+A=0\right\}
$$

(the set of all $n \times n$ skew-symmetric matrices in $\mathbb{R}^{n \times n}$ ).
On the other hand, if $A \in \operatorname{Sk}$-sym $(n) \subseteq \mathbb{R}^{n \times n}$, we consider the curve

$$
\alpha:(a, b) \rightarrow \mathrm{GL}(n, \mathbb{R}), \quad t \mapsto \exp (t A) .
$$

Then

$$
\begin{aligned}
\alpha(t)^{\top} \alpha(t) & =\exp (t A)^{\top} \exp (t A) \\
& =\exp \left(t A^{\top}\right) \exp (t A) \\
& =\exp (-t A) \exp (t A) \\
& =I .
\end{aligned}
$$

Hence we can view $\alpha$ as a curve in $\mathrm{O}(n)$. Since $\dot{\alpha}(0)=A$, this shows that

$$
\mathrm{Sk}-\operatorname{sym}(n) \subseteq T_{I} \mathrm{O}(n)
$$

and hence the Lie algebra $\mathfrak{o}(n)$ of the orthogonal group $\mathrm{O}(n)$ consists of all $n \times n$ skew-symmetric matrices, with the usual commutator as the Lie bracket. Thus

$$
\mathfrak{o}(n)=T_{I} \mathrm{O}(n)=\operatorname{Sk} \text {-sym }(n)=\left\{A \in \mathbb{R}^{n \times n} \mid A^{\top}+A=0\right\} .
$$

It follows that (see Proposition 4.4.9)

$$
\operatorname{dim} \mathrm{O}(n)=\operatorname{dim} \mathfrak{o}(n)=\frac{n(n-1)}{2} .
$$

$\diamond$ Exercise 238 Show that if $A \in \operatorname{Sk}$-sym ( $n$ ), then $\operatorname{tr} A=0$.
By Liouville's Formula, we have

$$
\operatorname{det} \alpha(t)=\operatorname{det} \exp (t A)=1
$$

and hence $\alpha:(a, b) \rightarrow \mathrm{SO}(n)$, where $\mathrm{SO}(n)$ is the special orthogonal group. We have actually shown that the Lie algebra of the special orthogonal group $\mathrm{SO}(n)$ is

$$
\mathfrak{s o}(n)=\mathfrak{o}(n)=\left\{A \in \mathbb{R}^{n \times n} \mid A^{\top}+A=0\right\} .
$$

## The Lie algebra of SO (3)

We will discuss the Lie algebra $\mathfrak{s o}(3)$ of the rotation group $\mathrm{SO}(3)$ in some detail.
$\diamond$ Exercise 239 Show that

$$
\mathfrak{s o}(3)=\left\{\left.\left[\begin{array}{ccc}
0 & -c & b \\
c & 0 & -a \\
-b & a & 0
\end{array}\right] \in \mathbb{R}^{3 \times 3} \right\rvert\, a, b, c \in \mathbb{R}\right\} .
$$

The Lie algebra $\mathfrak{s o}(3)$ is a 3 -dimensional real vector space. Consider the rotations
$R_{1}(t)=\left[\begin{array}{ccc}1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t\end{array}\right], R_{2}(t)=\left[\begin{array}{ccc}\cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t\end{array}\right], R_{3}=\left[\begin{array}{ccc}\cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1\end{array}\right]$.
Then the mappings

$$
\rho_{i}: t \mapsto R_{i}(t), \quad i=1,2,3
$$

are curves in $\mathrm{SO}(3)$ and clearly $\rho_{i}(0)=I$. It follows that

$$
\dot{\rho}_{i}(0):=A_{i} \in \mathfrak{s o}(3), \quad i=1,2,3 .
$$

These elements (matrices) are

$$
A_{1}=\left[\begin{array}{ccc}
0 & 0 & 0 \\
0 & 0 & -1 \\
0 & 1 & 0
\end{array}\right], \quad A_{2}=\left[\begin{array}{ccc}
0 & 0 & 1 \\
0 & 0 & 0 \\
-1 & 0 & 0
\end{array}\right], \quad A_{3}=\left[\begin{array}{ccc}
0 & -1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{array}\right]
$$

$\diamond$ Exercise 240 Verify that the matrices $A_{1}, A_{2}, A_{3}$ form a basis for $\mathfrak{s o}(3)$. We shall refer to this basis as the standard basis.
$\diamond$ Exercise 241 Compute all the Lie brackets (commutators) $\left[A_{i}, A_{j}\right], \quad i, j=$ $1,2,3$ and then determine the coefficients $c_{i j}^{k}$ defined by

$$
\left[A_{i}, A_{j}\right]=c_{i j}^{1} A_{1}+c_{i j}^{2} A_{2}+c_{i j}^{3} A_{3}, \quad i, j=1,2,3 .
$$

These coefficients are called the structure constants of the Lie algebra. They determine completely the Lie bracket $[\cdot, \cdot]$.

The Lie algebra $\mathfrak{s o}(3)$ may be identified with (the Lie algebra) $\mathbb{R}^{3}$ as follows. We define the mapping

$$
\mathfrak{\imath}: \mathbb{R}^{3} \rightarrow \mathfrak{s o}(3), \quad x=\left(x_{1}, x_{2}, x_{3}\right) \mapsto \widehat{x}:=\left[\begin{array}{ccc}
0 & -x_{3} & x_{2} \\
x_{3} & 0 & -x_{1} \\
-x_{2} & x_{1} & 0
\end{array}\right]
$$

This mapping is called the hat mapping.
$\diamond$ Exercise 242 Show that the hat mapping ${ }^{\wedge}: \mathbb{R}^{3} \rightarrow \mathfrak{s o}(3)$ is an isomorphism of vector spaces.
$\diamond$ Exercise 243 Show that (for $x, y \in \mathbb{R}^{3}$ )
(a) $x \times y=\widehat{x} y$.
(b) $\widehat{x \times y}=[\widehat{x}, \widehat{y}]$.
(c) $x \bullet y=-\frac{1}{2} \operatorname{tr}(\widehat{x} \widehat{y})$.

Formula (b) says that the hat mapping is in fact an isomorphism of Lie algebras and so we can identify the Lie algebra $\mathfrak{s o ( 3 )}$ with (the Lie algebra) $\mathbb{R}^{3}$.

Note : For $x \in \mathbb{R}^{3}$ and $t \in \mathbb{R}$, the matrix exponential $\exp (t \widehat{x})$ is a rotation about (the axis) $x$ through the angle $t\|x\|$. The following explicit formula for $\exp (\widehat{x})$ is known as Rodrigues' Formula :

$$
\exp (\widehat{x})=I+\frac{\sin \|x\|}{\|x\|} \widehat{x}+\frac{1}{2}\left[\frac{\sin \left(\frac{\|x\|}{2}\right)}{\frac{\|x\|}{2}}\right]^{2} \widehat{x}^{2} .
$$

This result says that the exponential mapping

$$
\exp : \mathfrak{s o}(3) \rightarrow \mathrm{SO}(3)
$$

is onto. Rodrigues' Formula is useful in computational solid mechanics, along with its quaternionic counterpart.

The Lie algebras of $\mathrm{U}(n)$ and $\mathrm{SU}(n)$
Consider the unitary group $\mathrm{U}(n)$; that is,

$$
\mathrm{U}(n)=\left\{A \in \mathrm{GL}(n, \mathbb{C}) \mid A^{*} A=I\right\} .
$$

For a curve $\alpha$ in $\mathrm{U}(n)$ with $\alpha(0)=I$, we obtain

$$
\dot{\alpha}(0)^{*}+\dot{\alpha}(0)=0
$$

and so $\dot{\alpha}(0) \in \mathbb{C}^{n \times n}$ is skew-Hermitian. So

$$
T_{I} \cup(n) \subseteq \operatorname{Sk} \text {-Herm }(n)=\left\{A \in \mathbb{C}^{n \times n} \mid A^{*}+A=0\right\}
$$

(the set of all $n \times n$ skew-Hermitian matrices in $\mathbb{C}^{n \times n}$ ).
If $H \in \operatorname{Sk}$ - $\operatorname{Herm}(n)$, then the curve

$$
\alpha:(a, b) \rightarrow \mathrm{GL}(n, \mathbb{C}), \quad t \mapsto \exp (t H)
$$

satisfies

$$
\begin{aligned}
\alpha(t)^{*} \alpha(t) & =\exp (t H)^{*} \exp (t H) \\
& =\exp \left(t H^{*}\right) \exp (t H) \\
& =\exp (-t H) \exp (t H) \\
& =I .
\end{aligned}
$$

Hence we can view $\alpha$ as a curve in $\mathrm{U}(n)$. Since $\dot{\alpha}(0)=H$, this shows that

$$
\text { Sk-Herm }(n) \subseteq T_{I} \cup(n)
$$

and hence the Lie algebra $\mathfrak{u}(n)$ of the unitary group $\mathrm{U}(n)$ consists of all $n \times n$ skew-Hermitian matrices, with the usual commutator as the Lie bracket. Thus

$$
\mathfrak{u}(n)=T_{I} \cup(n)=\text { Sk-Herm }(n)=\left\{H \in \mathbb{C}^{n \times n} \mid H^{*}+H=0\right\} .
$$

It follows that (see Exercise 229)

$$
\operatorname{dim} \mathcal{U}(n)=\operatorname{dim}_{\mathbb{R}} \mathfrak{u}(n)=n^{2} .
$$

The special unitary group $\operatorname{SU}(n)$ can be handled in a similar way. Again we have

$$
\mathfrak{s u}(n)=T_{I} \mathrm{SU}(n) \subseteq \operatorname{Sk} \text {-Herm }(n) .
$$

But also if $\alpha:(a, b) \rightarrow \mathrm{SU}(n)$ is a curve with $\alpha(0)=I$ then, as in the analysis for $\operatorname{SL}(n, \mathbb{R})$,

$$
\operatorname{tr} \dot{\alpha}(0)=0
$$

Writing

$$
\operatorname{Sk-Herm}^{0}(n):=\{H \in \operatorname{Sk}-\operatorname{Herm}(n) \mid \operatorname{tr} H=0\}
$$

this gives $\mathfrak{s u}(n) \subseteq \operatorname{Sk}_{\mathrm{k}}-\operatorname{Herm}^{0}(n)$. On the other hand, if $H \in \operatorname{Sk}^{\left(-\operatorname{Herm}^{0}\right.}(n)$ then the curve

$$
\alpha:(a, b) \rightarrow \mathrm{U}(n), \quad t \mapsto \exp (t H)
$$

takes values in $\mathrm{SU}(n)$ and $\dot{\alpha}(0)=H$. Hence
$\mathfrak{s u}(n)=T_{I} \mathrm{SU}(n)=\mathrm{Sk}_{\mathrm{k}}^{\mathrm{Herm}}{ }^{0}(n)=\left\{H \in \mathbb{C}^{n \times n} \mid H^{*}+H=0\right.$ and $\left.\operatorname{tr} H=0\right\}$.

Note : For a matrix group $G \leq \operatorname{GL}(n, \mathbb{R})$ (with Lie algebra $\mathfrak{g}$ ), the following are true (and can be used in determining Lie algebras of matrix groups).

- The mapping

$$
\exp _{G}: \mathfrak{g} \rightarrow \mathrm{GL}(n, \mathbb{R}), \quad X \mapsto \exp (X)
$$

has image contained in $G, \exp _{G}(\mathfrak{g}) \subseteq G$. We will normally write $\exp _{G}: \mathfrak{g} \rightarrow G$ for the exponential mapping on $G$ (and sometimes even just exp). In general, the exponential mapping $\exp _{G}$ is neither one-to-one nor onto.

- If $G$ is compact and connected, then $\exp _{G}$ is onto.
- The mapping $\exp _{G}$ maps a neighborhood of 0 in $\mathfrak{g}$ bijectively onto a neighborhood of $I$ in $G$.
$\diamond$ Exercise 244 Verify that the exponential mapping

$$
\exp _{(1)}: \mathbb{R} \rightarrow \mathbf{U}(1)=\mathbb{S}^{1}, \quad t \mapsto e^{i t}
$$

is onto but not one-to-one.
4.6.1 Example. The exponential mapping

$$
\exp _{(2, \mathbb{R})}: \mathfrak{s l}(2, \mathbb{R}) \rightarrow \operatorname{SL}(2, \mathbb{R})
$$

is not onto. Let

$$
A=\left[\begin{array}{ll}
\lambda & 0 \\
0 & \frac{1}{\lambda}
\end{array}\right] \quad \text { with } \lambda<-1 .
$$

We see that $A \in \mathrm{SL}(2, \mathbb{R})$ and we shall show that $A$ is not of the form $\exp (X)$ with $X \in \mathfrak{s l}(2, \mathbb{R})$. If $A=\exp (X)$, then the eigenvalues of $A$ are of the form $e^{a}$ and $e^{b}$, where $a$ and $b$ are the eigenvalues of $X$. Suppose $\lambda=e^{a}$ and $\frac{1}{\lambda}=e^{b}$. Then

$$
a=-b+2 k \pi i, \quad k \in \mathbb{Z} .
$$

However, since $\lambda$ is negative, $a$ is actually complex and therefore its conjugate is also an eigenvalue; that is, $b=\bar{a}$. This gives $a$ as pure imaginary. Then

$$
1=\left|e^{a}\right|=|\lambda|=-\lambda
$$

which contradicts the assumption that $\lambda<-1$.

## The Lie algebra of SU (2)

We will discuss the Lie algebra $\mathfrak{s u}(2)$ in some detail.
$\diamond$ Exercise 245 Show that

$$
\mathfrak{s u}(2)=\left\{\left.\left[\begin{array}{cc}
c i & -b+a i \\
b+a i & -c i
\end{array}\right] \in \mathbb{C}^{2 \times 2} \right\rvert\, a, b, c \in \mathbb{R}\right\} .
$$

The Lie algebra $\mathfrak{s u}(2)$ is a 3 -dimensional real vector space. Consider the matrices

$$
H_{1}=\frac{1}{2}\left[\begin{array}{ll}
0 & i \\
i & 0
\end{array}\right], \quad H_{2}=\frac{1}{2}\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right], \quad H_{3}=\frac{1}{2}\left[\begin{array}{cc}
i & 0 \\
0 & -i
\end{array}\right] .
$$

Clearly,

$$
H_{i} \in \mathfrak{s u}(2), \quad i=1,2,3 .
$$

Exercise 246 Verify that the matrices $H_{1}, H_{2}, H_{3}$ form a basis for $\mathfrak{s u}(2)$.
Exercise 247 Compute all the Lie brackets (commutators) $\left[H_{i}, H_{j}\right], i, j=$ $1,2,3$ and then determine the structure constants of (the Lie algebra) $\mathfrak{s u}(2)$.

Consider the mapping

$$
\phi: \mathbb{R}^{3} \rightarrow \mathfrak{s u}(2), \quad x=\left(x_{1}, x_{2}, x_{3}\right) \mapsto x_{1} H_{1}+x_{2} H_{2}+x_{3} H_{3} .
$$

Exercise 248 Show that the mapping $\phi: \mathbb{R}^{3} \rightarrow \mathfrak{s u}(2)$ is an isomorphism of Lie algebras.

Thus we can identify the Lie algebra $\mathfrak{s u}(2)$ with (the Lie algebra) $\mathbb{R}^{3}$.
Note : The Lie algebras $\mathfrak{s u ( 2 )}$ and $\mathfrak{s o}(3)$ look the same algebraically (they are isomorphic). An explicit isomorphism (of Lie algebras) is given by

$$
\psi: x_{1} H_{1}+x_{2} H_{2}+x_{3} H_{3} \mapsto x_{1} A_{1}+x_{2} A_{2}+x_{3} A_{3} .
$$

This suggests that there might be a close relationship between the matrix groups themselves. Indeed there is a (surjective) Lie homomorphism $\mathrm{SU}(2) \rightarrow \mathrm{SO}(3)$ whose derivative (at $I$ ) is $\psi$. Recall the adjoint representation

$$
\operatorname{Ad}: \operatorname{SU}(2) \rightarrow \operatorname{Aut}(\mathfrak{s u}(2)), \quad A \mapsto \operatorname{Ad}_{A}\left(: U \mapsto A U A^{*}\right) .
$$

Each $\mathrm{Ad}_{A}$ is a linear isomorphism of $\mathfrak{s u}(2) . \operatorname{Ad}_{A}$ is actually an orthogonal transformation on $\mathfrak{s u}(2)$ (the mapping $(X, Y) \mapsto-\operatorname{tr}(X Y)$ is an inner product on $\mathfrak{s u}(2))$ and so $\mathrm{Ad}_{A}$ corresponds to an element of $\mathrm{O}(3)$ (in fact, $\mathrm{SO}(3)$ ). The mapping

$$
\overline{\mathrm{Ad}}: \mathrm{SU}(2) \rightarrow \mathrm{SO}(3), \quad A \mapsto \mathrm{Ad}_{A}
$$

turns out to be a continuous homomorphism of matrix groups that is differentiable (i.e., a Lie homomorphism) and such that its derivative $d \overline{\mathrm{Ad}}: \mathfrak{s u}(2) \rightarrow \mathfrak{s o}(3)$ is $\psi$.

## The Lie algebras of $\mathrm{UT}(n, \mathbb{k})$ and $\mathrm{UT}^{u}(n, \mathbb{k})$

Let $\alpha:(a, b) \rightarrow \mathrm{UT}(n, \mathbb{k})$ be a curve in $\mathrm{UT}(n, \mathbb{k})$ with $\alpha(0)=I$. Then $\dot{\alpha}(0)$ is upper triangular. Moreover, using the argument for $\mathrm{GL}(n, \mathbb{k})$ we see that given any upper triangular matrix $A \in \mathbb{k}^{n \times n}$, there is a curve

$$
\sigma:(-\epsilon, \epsilon) \rightarrow \mathbb{k}^{n \times n}, \quad t \mapsto t A+I
$$

such that $\sigma(t) \in \mathrm{UT}(n, \mathbb{k})$ and $\dot{\sigma}(0)=A$. Hence the Lie algebra $\mathfrak{u t}(n, \mathbb{k})$ of $\mathrm{UT}(n, \mathbb{k})$ consists of all $n \times n$ upper triangular matrices, with the usual commutator as the Lie bracket. Thus

$$
\mathfrak{u t}(n, \mathbb{k})=T_{I} \cup \mathbf{\top}(n, \mathbb{k})=\left\{A \in \mathbb{k}^{n \times n} \mid a_{i j}=0 \text { for } i>j\right\} .
$$

It follows that

$$
\operatorname{dim} U T(n, \mathbb{k})=\operatorname{dim}_{\mathbb{R}} \mathfrak{u t}(n, \mathbb{k})=\frac{n(n+1)}{2} \operatorname{dim}_{\mathbb{R}} \mathbb{k}
$$

An upper triangular matrix $A \in \mathbb{k}^{n \times n}$ is strictly upper triangular if all its diagonal entries are 0 . Then the Lie algebra of the unipotent group $\mathrm{UT}^{u}(n, \mathbb{k})$ consists of all $n \times n$ strictly upper triangular matrices, with the usual commutator as the Lie bracket. So

$$
\mathfrak{s u t}(n, \mathbb{k})=T_{I} \cup^{u}(n, \mathbb{k})=\left\{A \in \mathbb{k}^{n \times n} \mid a_{i j}=0 \text { for } i \geq j\right\}
$$

$\diamond$ Exercise 249 Find $\operatorname{dim}_{\mathbb{R}} \mathfrak{s u t}(n, \mathbb{k})$.
$\diamond$ Exercise 250 For each of the following matrix group $G$, determine its Lie algebra $\mathfrak{g}$ and hence its dimension.
(a) $G=\left\{A \in \mathrm{GL}(2, \mathbb{R}) \mid A Q A^{\top}=Q\right\}$, where $Q=\left[\begin{array}{ll}1 & 0 \\ 0 & 0\end{array}\right]$.
(b) $G=\left\{A \in \mathrm{GL}(2, \mathbb{R}) \mid A Q A^{\top}=Q\right\}$, where $Q=\left[\begin{array}{cc}1 & 0 \\ 0 & -1\end{array}\right]$.
(c) $G=\mathrm{GA}(3, \mathbb{R})$.
(d) $G=$ Heis.
(e) $G=G_{4} \leq \mathrm{UT}^{u}(4, \mathbb{R})$ from Exercise 199.
(f) $G=\mathrm{E}(n)$.
(g) $G=\operatorname{SE}(n)$.
$\diamond$ Exercise 251
(a) Show that the Lie algebra of the symplectic group $\operatorname{Sp}(2 n, \mathbb{R})$ is

$$
\mathfrak{s p}(2 n, \mathbb{R})=\left\{A \in \mathbb{R}^{2 n \times 2 n} \mid A^{\top} \mathbb{J}+\mathbb{J} A=0\right\} .
$$

(b) If

$$
A=\left[\begin{array}{ll}
a & b \\
c & d
\end{array}\right] \in \mathfrak{s l}(2 n, \mathbb{R})
$$

show that $A \in \mathfrak{s p}(2 n, \mathbb{R})$ if and only if

$$
d=-a^{\top}, \quad c=c^{\top}, \quad \text { and } \quad b=b^{\top} .
$$

(c) Calculate the dimension of $\mathfrak{s p}(2 n, \mathbb{R})$.

Exercise 252 Show that the Lie algebra of the Lorentz group Lor is
$\mathfrak{l o r}=\left\{A \in \mathbb{R}^{4 \times 4} \mid S A+A^{\top} S=0\right\}=\left\{\left.\left[\begin{array}{cccc}0 & a_{1} & a_{2} & a_{3} \\ -a_{1} & 0 & a_{4} & a_{5} \\ -a_{2} & -a_{4} & 0 & a_{6} \\ a_{3} & a_{5} & a_{6} & 0\end{array}\right] \right\rvert\, a_{1}, a_{2}, a_{3}, a_{4}, a_{5}, a_{6} \in \mathbb{R}\right\}$.
$\diamond$ Exercise 253 Consider the matrix group $\mathbb{k}^{\times}=G L(1, \mathbb{k})$. (Its Lie algebra is clearly $\mathbb{k}$.)
(a) Show that the determinant function

$$
\operatorname{det}: \operatorname{GL}(n, \mathbb{k}) \rightarrow \mathbb{k}^{\times}
$$

is a Lie homomorphism (i.e. a continuous homomorphism of matrix groups that is also differentiable; cf. Definition 4.4.14).
(b) Show that the induced homomorphism of Lie algebras (i.e. the derivative of det) is the trace function

$$
\operatorname{tr}: \mathbb{K}^{n \times n} \rightarrow \mathbb{k} .
$$

(c) Derive from (b) that (for $A, B \in \mathbb{K}^{n \times n}$ )

$$
\operatorname{tr}(A B)=\operatorname{tr}(B A)
$$

