# Group Theoretical Aspects of Quantum Mechanics 

José Alvarado

December 4, 2007

## Contents

1 Groups ..... 1
1.1 Definition ..... 1
Multiplication Tables ..... 2
Cyclic Groups ..... 3
Dihedral Groups ..... 5
1.2 Subgroups ..... 7
Definition ..... 7
Conjugate Elements ..... 7
Normal Subgroups ..... 8
Cosets ..... 8
Factor Groups ..... 9
1.3 Homomorphisms ..... 12
Definition ..... 12
Subgroups Associated with Homomorphisms ..... 13
Isomorphisms ..... 13
The First Isomorphism Theorem ..... 15
1.4 Algebras ..... 16
Definition ..... 16
Structure Constants ..... 16
Lie Algebras ..... 17
Homomorphisms Between Algebras ..... 18
2 Representations ..... 19
2.1 General Representations ..... 19
The General Linear Group $\mathrm{GL}(V)$ ..... 19
Representations ..... 20
Terminology ..... 21
Invariance and Reducibility ..... 22
2.2 Matrix Representations ..... 22
The General Linear Group and its Subgroups in Matrix Form ..... 23
Terminology ..... 23
Characters and Equivalent Representations ..... 24
Reducibility ..... 24
2.3 Irreducible Representations ..... 26
3 Rotations ..... 31
3.1 The Lie Group SO(2) ..... 31
Matrix Representations of $S O(2)$ ..... 32
Manifolds and Lie groups ..... 33
Infinitesimal Generators ..... 33
Lie Algebras ..... 33
3.2 The Lie Group SO(3) ..... 34
The Lie algebra so(3) ..... 34
Euler-Angle Parametrization ..... 35
Axis-Angle Parametrization ..... 36
Topology of SO(3) ..... 36
Representation of SO(3) and so(3) on Wave Functions ..... 37
3.3 The Lie Group $\operatorname{SU}(2)$ ..... 38
The Lie algebra su(2) ..... 39
The Vector space $\mathrm{H}_{2}$ ..... 40
Rotating vectors in $H_{2}$ and $\mathbb{R}^{3}$ ..... 40
Axis-Angle Parametrization ..... 41
Lie Group Homomorphism $\operatorname{SU}(2) \rightarrow \mathrm{SO}(3)$ ..... 41
Lie Algebra Isomorphism su(2) $\rightarrow$ so(3) ..... 43
Representations of SO(3) and SU(2) ..... 43
4 Angular Momentum ..... 45
4.1 Spin Angular Momentum ..... 46
The Lie-Algebra $\mathbf{s l}(2, \mathbb{C})$ ..... 46
The Casimir Operator $\hat{\mathbf{J}}^{2}$ ..... 46
The Eigenvalue Ladder ..... 47
The irreducible representations of su(2) ..... 50
4.2 Orbital Angular Momentum and Rotational Symmetry ..... 53
Step 1: From su(2) to $\mathrm{SU}(2)$ ..... 53
Step 2: From SU(2) to SO(3) ..... 54
Basis of Irreducible Functions of $\mathrm{SO}(3)$ ..... 55
Partial Wave Decomposition ..... 57
Spherically Symmetric Potentials ..... 58
5 Adding Angular Momenta ..... 59
Tensor Product of Lie-Group Representations ..... 59
Clebsch-Gordan Series for $\operatorname{SU}(2)$ and $\mathrm{SO}(3)$ ..... 59
Tensor Product of Lie-Algebra Representations ..... 60
Clebsch-Gordan Coefficients ..... 60
Example: Two Particles of spin $\frac{1}{2}$ ..... 61
6 The Hidden Symmetry of the Hydrogen Atom ..... 64
The Quantum Number $n$ ..... 65
The Quantum Number $l$ ..... 66

## Chapter 1

## Groups

The easiest way to understand the concept of a group is to consider a simple example: the group $\mathrm{C}_{4}$. It consists of the numbers $i,-1,-i, 1$, which have the property that:

$$
\mathrm{i}^{1}=\mathrm{i} \quad \mathrm{i}^{2}=-1 \quad \mathrm{i}^{3}=-\mathrm{i} \quad \mathrm{i}^{4}=1
$$

Further powers of i will repeat this cycle, so $\mathrm{i}^{5}=\mathrm{i}$, and so on. In fact, this group is closed; multiplying any of these two numbers will always give us another one of these numbers. This is the kind of structure that separates groups from ordinary sets.

### 1.1 Definition

Formally, a set $G=\{a, b, c, \ldots\}$ forms a group $G$ if the following three conditions are satisfied:

- The set is accompanied by some operation • called group multiplication, which is associative: $a \cdot(b \cdot c)=(a \cdot b) \cdot c$.
- A special element $\mathbb{1} \in \mathrm{G}$, called the identity, has the property $a \cdot \mathbb{1}=\mathbb{1} \cdot a=a$ for all $a \in \mathrm{G}$.
- Each $a$ in G has a unique inverse $a^{-1}$ which must also be in G and satisfy $a \cdot a^{-1}=$ $a^{-1} \cdot a=\mathbb{1}$.

The group $C_{4}$ mentioned above satisfies these three conditions with standard numeric multiplication as group multiplication, and with the number 1 as the identity element $\mathbb{1}$. The inverses are:

$$
\mathrm{i}^{-1}=-\mathrm{i} \quad-1^{-1}=-1 \quad-\mathrm{i}^{-1}=\mathrm{i}
$$

Group multiplication isn't restricted to regular numeric multiplication. For instance, the set of all integers $\mathbb{Z}=\{\ldots,-2,-1,0,1,2, \ldots\}$ forms a group under standard algebraic addition + , and with the number 0 as the identity element $\mathbb{1}$. Every element has a unique
inverse in $\mathbb{Z}$ under addition; for instance, $3^{-1}=-3$. Notice that numeric multiplication wouldn't work, because $\frac{1}{3}$ is not in $\mathbb{Z}$.
The symbol ${ }^{-1}$ does not always mean "raise to the -1 th power," rather it means "find the inverse under the specified group multiplication." The same applies for positive superscripts, which mean "apply the specified group multiplication $n$ times." For instance, under addition, $4^{2}=8$.
A group is called abelian if the group multiplication - is commutative: $a \cdot b=b \cdot a$. $\mathrm{C}_{4}$ is clearly abelian, since all numbers commute under multiplication. An example of a non-abelian group is a set of matrices under matrix multiplication.

## Multiplication Tables

The order of a group is the number of elements it contains. For example, $\mathrm{C}_{4}$ is of order 4 , and $\mathbb{Z}$ is of infinite order.
There are two kinds of groups of infinite order: countable and continuous. The first kind has elements that can be counted, much like the integers $\mathbb{Z}$. The second kind cannot be counted, much like the real numbers $\mathbb{R}$. Continuous groups will lead us to the famous Lie Groups, which we will discuss later.
Groups of finite order can be summarized in convenient multiplication tables, which demonstrate how the group multiplication works on its elements. The following is a multiplication table for $\mathrm{C}_{4}$ :

| $\mathrm{C}_{4}$ | $\mathbb{1}$ | $a$ | $a^{2}$ | $a^{3}$ |
| :---: | :---: | :---: | :---: | :---: |
| $\mathbb{1}$ | $\mathbb{1}$ | $a$ | $a^{2}$ | $a^{3}$ |
| $a$ | $a$ | $a^{2}$ | $a^{3}$ | $\mathbb{1}$ |
| $a^{2}$ | $a^{2}$ | $a^{3}$ | $\mathbb{1}$ | $a$ |
| $a^{3}$ | $a^{3}$ | $\mathbb{1}$ | $a$ | $a^{2}$ |

Figure 1.1: Multiplication Table for $\mathrm{C}_{4}$
We use a more general notation, where we use $a$ instead of i .
Notice how each entry in the multiplication table is also a member of the group. This is true for all groups, and it illustrates their closure: if $a$ and $b$ are in G, then so is $a \cdot b$. This is often referred to as a fourth defining property of a group. (It's technically superfluous, since this property is already included in the mathematical definition of an operation. However, it's an easy property to overlook, so we mention it for emphasis.)
Notice also how $\mathbb{1}$ appears exactly once in every row and column. This is a consequence of the fact that every group element must have a unique inverse belonging to the group. Because of this, every group element appears only once in every row and column of a multiplication table. This kind of structure is called a Latin square.
Notice also that the multiplication table for $\mathrm{C}_{4}$ can be reflected along its main diagonal. This indicates that the group is abelian, that is, the order of multiplication doesn't matter: $a \cdot b=b \cdot a$

| $\mathrm{C}_{4}$ | $\mathbb{1}$ | $a$ | $a^{2}$ | $a^{3}$ |
| :---: | :---: | :---: | :---: | :---: |
| $\mathbb{1}$ | $\mathbb{1}$ | $a$ | $a^{2}$ | $a^{3}$ |
| $a$ | $\cdot$ | $a^{2}$ | $a^{3}$ | $\mathbb{1}$ |
| $a^{2}$ | $\cdot$ | $\cdot$ | $\mathbb{1}$ | $a$ |
| $a^{3}$ | $\cdot$ | $\cdot$ | $\cdot$ | $a^{2}$ |

Figure 1.2: We know what goes in place of the dots, because the group is abelian.

## Cyclic Groups

So far, we have covered many important properties of groups. But why should we even care about them? Why are they useful?
The short answer is: symmetry. Group theory provides us with a detailed, mathematical description of symmetry. Several problems in modern physics exhibit certain kinds of symmetries we'd like to take advantage of. Some are easy to grasp, whereas some are not so easily apparent. For instance, we will see at the end of this script that the hydrogen atom corresponds to the symmetry group of four-dimensional rotations. This kind of symmetry may be impossible to visualize, but it's still possible to study using group theory.
Although we're not ready to tackle a four-dimensional rotation group yet, we'll start out with a simple example which can be readily visualized. The cyclic group $C_{n}$ is defined to be the group of order $n$ with the following structure:

$$
\mathrm{C}_{n}=\left\{\mathbb{1}, a, a^{2}, a^{3}, \ldots, a^{n-1}\right\} \quad a^{n}=\mathbb{1}
$$

Notice that the entire group is determined by a single element $a$; we call this element the generator of the group. In order for $a$ to generate a cyclic group, it must satisfy $a^{n}=\mathbb{1}$ for some $n$. We are already familiar with the group $C_{4}$, which is the cyclic group of order 4. In fact, the name cyclic group lends its name from the rows and columns of its multiplication table, which are just cyclic permutations of each other.
This seemingly formal mathematical description of the cyclic group harbors geometrical significance: $\mathrm{C}_{n}$ corresponds to the symmetry of an $n$-sided polygon. (Actually, the cyclic group corresponds to the symmetry of an oriented polygon. This will soon become clear.) In order to investigate this claim, let's consider the group $C_{4}$. We should expect $C_{4}$ to somehow relate to a 4 -sided polygon, i.e. a square.
Let us consider an actual square, and rotate it as in Figure 1.3. We would certainly notice a rotation by a small angle, say $5^{\circ}$. However, if we were to rotate it by exactly $90^{\circ}$, we wouldn't be able to distinguish it before or after rotation. We can again rotate the square to $180^{\circ}, 270^{\circ}$, and back to $360^{\circ}$, and the square will still appear unchanged.
This property must somehow relate to the set $\left\{a, a^{2}, a^{3}, \mathbb{1}\right\}$. This can be acheived if we allow $a$ to be defined as "a rotation by $90^{\circ}$," as strange as that may seem. It would then follow that $a^{2}$ would mean two rotations by $90^{\circ}, a^{3}$ would mean three rotations by $90^{\circ}$, and $a^{4}=\mathbb{1}$ would mean four rotations by $90^{\circ}$ (i.e. do nothing). This interpretation for $a$ is commonly used in molecular physics to describe molecules with symmetry.
But what about the set $\{i,-1,-i, 1\}$ ? How does this set geometrically relate to $C_{4}$ ? If we connect these four points on the complex plane, we immediately notice a square tilted


Figure 1.3: Rotating a square by $90^{\circ}$ leaves it unchanged.
to look like a diamond.


Figure 1.4: $\mathrm{C}_{4}$ symmetry in the complex plane
Notice that multiplication by $e^{\mathrm{i} \frac{\pi}{2}}=\mathrm{i}$ will rotate a complex number by $\frac{\pi}{2}=90^{\circ}$, so our previous interpretation of $a$ holds in the complex plane with $a=i$.
We can rewrite the elements of $\mathrm{C}_{4}$ in the following form:

$$
\begin{gathered}
i=e^{i \frac{\pi}{2}}=e^{\frac{1}{4} \cdot 2 \pi i} \quad-1=e^{i \pi}=e^{\frac{2}{4} \cdot 2 \pi i} \quad-i=e^{i \frac{3 \pi}{2}}=e^{\frac{3}{4} \cdot 2 \pi i} \quad 1=e^{2 \pi i}=e^{\frac{4}{4} \cdot 2 \pi i} \\
\mathrm{C}_{4}=\left\{e^{\frac{1}{4} 2 \pi i}, e^{\frac{2}{4} 2 \pi i}, e^{\frac{3}{4} 2 \pi i}, e^{\frac{4}{4} 2 \pi i}\right\} .
\end{gathered}
$$

Therefore, another way of expressing the cyclic group $C_{4}$ is to consider the generator:

$$
a=e^{\frac{1}{4} \cdot 2 \pi i}=e^{\frac{2 \pi i}{4}} .
$$

We can easily extend this idea to cyclic groups of all orders $n$ by slightly modifying the above formula:

$$
a=e^{\frac{2 \pi i}{n}} .
$$

This generator exponentiates to give us all the elements of the cyclic group, including

$$
a^{n}=\left(e^{\frac{2 \pi i}{n}}\right)^{n}=e^{2 \pi i}=1=\mathbb{1} .
$$

We see in Figure 1.5 that the group elements form $n$-sided polygons in the complex plane.


Figure 1.5: $n$-sided polygons in the complex plane for $n=3,4,5$

## Dihedral Groups

Until now, we've covered the rotational symmetry of a square using the cyclic group $\mathrm{C}_{4}$. However, squares have another kind of symmetry: reflections. This kind of symmetry can be described using the dihedral group $\mathrm{D}_{n}$ of order $2 n$, which has the following structure:

$$
\mathrm{D}_{n}=\left\{\mathbb{1}, a, a^{2}, \ldots, a^{n-1}, b, b a, b a^{2}, \ldots, b a^{n-1}\right\} \quad a^{n}=b^{2}=(a b)^{2}=\mathbb{1}
$$

Notice that the dihedral group requires two generators $a$ and $b$, which must satisfy the three equations given above. As in the cyclic group, $a$ retains its interpretation as a rotation. Our new generator $b$, however, corresponds to a reflection.
In order to fully show the difference between rotations and reflections, we need to consider an oriented polygon: take a square, and add arrows to its sides.


Figure 1.6: An oriented square
Rotating this oriented square by $90^{\circ}$ will give us the same square with the same orientation again. However, if we reflect it across the $y$-axis, we get the same square but not the same orientation.
The group $D_{4}$ is a group of order eight, with elements

$$
\mathrm{D}_{4}=\left\{\mathbb{1}, a, a^{2}, a^{3}, b, b a, b a^{2}, b a^{3}\right\}
$$

and the defining identities

$$
a^{4}=b^{2}=(b a)^{2}=\mathbb{1}
$$

This definition of $\mathrm{D}_{4}$ is indeed complete. Other combinations of $a$ and $b$ are possible, but they always reduce to one of the eight elements given above. For instance, we have

$$
a b=b a^{3} .
$$



Figure 1.7: Rotation and reflection of an oriented square

This shows us that the dihedral groups are clearly non-abelian. You can test this identity visually by drawing arrows on a square piece of paper, making it oriented. Also draw an axis through the middle around which to flip, as well as a dot in one of the corners to keep track of the square's motion.


Figure 1.8: Decorate a piece of paper like this to understand $D_{4}$.
If you rotate the paper counter-clockwise by $90^{\circ}$ and then flip it ( $a b$ ), you'll get the same result by first flipping and then rotating by $270^{\circ}\left(b a^{3}\right)$. Be sure to focus on the dot's initial and final position.
We can prove this in a more mathematical fashion. If we take the identity

$$
b a b a=\mathbb{1},
$$

multiply by $b$ on the left and $a^{3}$ on the right, we instantly get our desired identity:

$$
\begin{aligned}
b^{2} a b a^{4} & =b \mathbb{1} a^{3} \\
a b & =b a^{3} .
\end{aligned}
$$

A good exercise for the reader is to prove the following identities, both visually and mathematically:

$$
a^{2} b=b a^{2} \quad a^{3} b=b a .
$$

### 1.2 Subgroups

## Definition

A subgroup $S$ of a group $G$ is defined to contain a subset of $G$ and satisfy the three defining conditions of a group using the same kind of group multiplication. For instance, the group $C_{4}=\left\{\mathbb{1}, a, a^{2}, a^{3}\right\}$ is a subgroup of $\mathrm{D}_{4}=\left\{\mathbb{1}, a, a^{2}, a^{3}, b, b a, b a^{2}, b a^{3}\right\}$.
$\mathrm{D}_{4}$ has two more subgroups. One is $\mathrm{D}_{4}$ itself, since for any set $S$ we have $S \subset S$. Another subgroup is simply $\{\mathbb{1}\}$. In fact, every group has these two subgroups: itself, and the identity. These are known as the trivial subgroups. We're usually interested in nontrivial subgroups, which are sometimes referred to as proper subgroups.

## Conjugate Elements

Two elements $a$ and $b$ of a group $G$ are said to be conjugate if there exists another element $g \in \mathrm{G}$, called the conjugating element, so that

$$
a=g b g^{-1} .
$$

For instance, the elements $a$ and $a^{3}$ of the group $\mathrm{D}_{4}$ are conjugate since

$$
b a b^{-1}=b a b=b b a^{3}=a^{3} .
$$

We denote this conjugacy relation by $a \sim a^{3}$. In fact, conjugation is an example of an equivalence relation, i.e., a relation with the following properties:

$$
\begin{aligned}
\text { Reflexivity: } & a \sim a \\
\text { Symmetry: } & a \sim b \Rightarrow b \sim a \\
\text { Transitivity: } & a \sim b \wedge b \sim c \Rightarrow a \sim c
\end{aligned}
$$

Group elements that are conjugate to each other are also said to be similar. In the case of $a \sim a^{3}$, we can actually visualize this: a $90^{\circ}$ rotation about one axis is the same as a $270^{\circ}$ rotation about the opposite axis. The elements $a$ and $a^{3}$ are certainly not equal, but at least they are similar.
One important detail to keep in mind is that the conjugating element $g$ must be a member of the group G. For instance, if we consider the group $\mathrm{C}_{4}$, the relation $a \sim a^{3}$ does not hold anymore, since the conjugating element $b$ does not belong to $C_{4}$.
We can list all the conjugation relations of the group $\mathrm{D}_{4}$ :

$$
\mathbb{1} \quad a \sim a^{3} \quad a^{2} \quad b \sim b a^{2} \quad b a \sim b a^{3} .
$$

Notice that we've neatly organized all eight elements of this group according to conjugacy. We can thus introduce the notation

$$
\mathrm{D}_{4}=\left\{(\mathbb{1}),(a),\left(a^{2}\right),(b),(b a)\right\},
$$

where the parentheses denote a conjugacy class, defined by the set of all conjugate elements:

$$
(a)=\{g \in \mathbf{G} \mid g \sim a\} .
$$

Since conjugation is an equivalence relation, these conjugacy classes cannot overlap each other; each element belongs to only one conjugacy class.

## Normal Subgroups

A subgroup N of G is called a normal subgroup if

$$
g n g^{-1} \in \mathbf{N} \quad \forall n \in \mathbf{N} \forall g \in \mathbf{G}
$$

That is, $\mathbf{N}$ isolates conjugate elements from the rest of the group so that no element $n \in \mathbf{N}$ can be conjugate to another foreign element $g \notin \mathrm{~N}$. We denote a normal subgroup by $N \triangleleft G$.
A normal subgroup always consists of a combination of complete conjugacy classes, so they are easy to spot if the parent group G is broken down into conjugacy classes. For instance, $C_{4}$ is a normal subgroup of $D_{4}$, since $C_{4}$ consists of three conjugacy classes:

$$
\left\{(\mathbb{1}),(a),\left(a^{2}\right)\right\}=\left\{\mathbb{1}, a, a^{2}, a^{3}\right\}=\mathrm{C}_{4} \triangleleft \mathrm{D}_{4} .
$$

Another possible normal subgroup of $D_{4}$ is

$$
\left\{(\mathbb{1}),\left(a^{2}\right)\right\}=\left\{\mathbb{1}, a^{2}\right\}=\mathrm{C}_{2} \triangleleft \mathrm{D}_{4} .
$$

Any combination of conjugacy classes forms a normal subgroup, as long as the group properties are fulfilled. For instance, the set

$$
\{(\mathbb{1}),(a)\}=\left\{\mathbb{1}, a, a^{3}\right\}
$$

may consist of complete conjugacy classes, but it does not form a group since it is not closed-it does not contain $a \cdot a$.
As another example, let's consider the group $\mathbb{R}$ of real numbers under addition. We see that no two group elements $x \in \mathbb{R}$ are conjugate, since

$$
y \cdot x \cdot y^{-1}=y+x-y=x \quad \forall y \in \mathbb{R}
$$

(In fact, this is true for any Abelian group.) Since every number makes up its own conjugacy class, any subgroup of $\mathbb{R}$ is a normal subgroup. For instance, we take the group $\mathbb{Z}$ of integers under addition, which is a normal subgroup of $\mathbb{R}$ :

$$
\mathbb{Z} \triangleleft \mathbb{R}
$$

## Cosets

If we have a subgroup $\mathbf{S}=\left\{s_{1}, s_{2}, \ldots\right\}$ of a group $G$, we can form an object $g$ S, called a coset, which is formed by premultiplying all elements of $S$ with $g$ :

$$
g S:=\left\{g s_{1}, g s_{2}, \ldots\right\}
$$

For instance, if we consider the subgroup $\mathrm{C}_{4}$ of $\mathrm{D}_{4}$, we can build the coset

$$
a \mathrm{C}_{4}=\left\{a, a^{2}, a^{3}, \mathbb{1}\right\}=\mathrm{C}_{4}
$$

Since we're not too worried about the order of elements within a set, we can safely say that $a \mathrm{C}_{4}=\mathrm{C}_{4}$.

To be precise, we have just defined a left coset. A right coset $S g$ is formed by postmultiplying all elements of $S$ with $g$ :

$$
\mathrm{S} g:=\left\{s_{1} g, s_{2} g, \ldots\right\}
$$

Using cosets, we can provide an equivalent definition of a normal subgroup:

$$
\mathrm{N} \triangleleft \mathrm{G} \quad: \Leftrightarrow g \mathrm{~N} g^{-1}=\mathrm{N} \quad \forall g \in \mathrm{G} .
$$

Postmultiplying this equation by $g$ shows us that for a normal subgroup, all left and right cosets are equal:

$$
g \mathbf{N}=\mathbf{N} g \quad \forall g \in \mathbf{G} .
$$

## Factor Groups

Given a group G and a normal subgroup N, we define the factor group as the set of all cosets:

$$
\mathrm{G} / \mathrm{N}:=\{g \mathrm{~N}\}_{g \in G} .
$$

For instance, the factor group $D_{4} / C_{4}$ is easy to calculate:

$$
\begin{array}{r}
\mathbb{1 C}_{4}=\left\{\mathbb{1}, a, a^{2}, a^{3}\right\}=\mathrm{C}_{4} \equiv E \\
a \mathrm{C}_{4}=\left\{a, a^{2}, a^{3}, \mathbb{1}\right\}=\mathrm{C}_{4} \equiv E \\
a^{2} \mathrm{C}_{4}=\left\{a^{2}, a^{3}, \mathbb{1}, a\right\}=\mathrm{C}_{4} \equiv E \\
a^{3} \mathrm{C}_{4}=\left\{a^{3}, \mathbb{1}, a, a^{2}\right\}=\mathrm{C}_{4} \equiv E \\
b \mathrm{C}_{4}=\left\{b, b a, b a^{2}, b a^{3}\right\}=b \mathrm{C}_{4} \equiv A \\
b a \mathrm{C}_{4}=\left\{b a, b a^{2}, b a^{3}, b\right\}=b \mathrm{C}_{4} \equiv A \\
b a^{2} \mathrm{C}_{4}=\left\{b a^{2}, b a^{3}, b, b a\right\}=b \mathrm{C}_{4} \equiv A \\
b a^{3} \mathrm{C}_{4}=\left\{b a^{3}, b, b a, b a^{2}\right\}=b \mathrm{C}_{4} \equiv A .
\end{array}
$$

Using the notation $\mathrm{C}_{4} \equiv E$ and $b \mathrm{C}_{4} \equiv A$, we have

$$
\mathrm{D}_{4} / \mathrm{C}_{4}=\{E, A\}
$$

So far, this seems to be a pointless set ornamented with arbitrary notation. However, $\{E, A\}$ is more than just a set. If we define an operation $\star$ among cosets in the following way,

$$
(g N) \star(h N):=(g \cdot h) N,
$$

and if we notice that this operation fulfills the defining properties of a group, we see that our set of cosets actually forms a group under the operation defined above. (In fact, we could always define a factor set G/S for a non-normal subgroup S. However, this set does not form a group. See Jones, §2.3.) Now that we know that $\mathrm{D}_{4} / \mathrm{C}_{4}$ forms a group, all that's left to do is to find out what kind of group it is. If we notice that

$$
\begin{aligned}
E \star E & =\left(\mathbb{1} \mathrm{C}_{4}\right) \star\left(\mathbb{1} \mathrm{C}_{4}\right)=\mathbb{1} \mathrm{C}_{4}=E \\
A \star E=E \star A & =\left(\mathbb{1} \mathrm{C}_{4}\right) \star\left(a \mathrm{C}_{4}\right)=a \mathrm{C}_{4}=A \\
A \star A & =\left(a \mathrm{C}_{4}\right) \star\left(a \mathrm{C}_{4}\right)=\mathbb{1} \mathrm{C}_{4}=E,
\end{aligned}
$$

we see that the group

$$
\{E, A\} \quad A^{2}=E
$$

behaves exactly like the cyclic group of order two:

$$
\mathrm{C}_{2}=\{\mathbb{1}, a\} \quad \quad a^{2}=\mathbb{1}
$$

Thus, we have

$$
\mathrm{D}_{4} / \mathrm{C}_{4}=\mathrm{C}_{2} .
$$

(To be correct, we should write $\mathrm{D}_{4} / \mathrm{C}_{4} \cong \mathrm{C}_{2}$, that is, the two groups are isomorphic. We'll define this term in the next section.)
A good exercise for the reader is to show that the following relations hold:

$$
\mathrm{D}_{4} / \mathrm{C}_{2}=\mathrm{D}_{2} \quad \mathrm{D}_{4} / \mathrm{D}_{2}=\mathrm{C}_{2}
$$

where

$$
\mathrm{D}_{2}=\{\mathbb{1}, a, b, b a\} \quad a^{2}=b^{2}=(b a)^{2}=\mathbb{1}
$$

Be sure to keep track of your $a$ 's. The element $a \in \mathrm{C}_{2}, \mathrm{D}_{2}$ corresponds to a $180^{\circ}$ rotation and is worth two elements $a \in \mathrm{C}_{4}, \mathrm{D}_{4}$, which correspond to a $90^{\circ}$ rotation.
Intuitively, the factor group $\mathrm{G} / \mathrm{N}$ treats the normal subgroup N as an identity element $E=\mathbb{1} \mathrm{N}$. We should somehow expect all the cosets to "collapse" under this new identity. Recall that $\mathrm{D}_{4} / \mathrm{C}_{4}$ had eight possible cosets, but they were all equal to either $E$ or $A$. We then had to determine the structure of this collapsed group.
Let's explore another example. Recall that we recently showed that $\mathbb{Z} \triangleleft \mathbb{R}$. What kind of group should we expect from $\mathbb{R} / \mathbb{Z}$ ?
We'll start by looking at the set of cosets $x+\mathbb{Z}$ :

$$
x+\mathbb{Z}=x+\{\ldots,-1,0,1,2, \ldots\}=\{\ldots,-1+x, x, 1+x, 2+x, \ldots\}
$$

We first look at the new identity

$$
0+\mathbb{Z}=\{\ldots,-1,0,1,2, \ldots\}=\mathbb{Z} \equiv E
$$

If we consider cosets of small $x$, we simply end up with $x+\mathbb{Z}=E+x$ :

$$
\begin{aligned}
0.1+\mathbb{Z} & =\{\ldots,-0.9,0.1,1.1,2.1, \ldots\} \\
0.2+\mathbb{Z} & =\{\ldots,-0.8,0.2,1.2,2.2, \ldots\}
\end{aligned}
$$

However, once we reach $x=1$, we get $1+\mathbb{Z}=\mathbb{Z} \equiv E$ :

$$
\begin{aligned}
& 0.9+\mathbb{Z}=\{\ldots,-0.1,0.9,1.9,2.9, \ldots\} \equiv E+0.9 \\
& 1+\mathbb{Z}=\{\ldots,-0,1,2,3, \ldots\}=\mathbb{Z}=\equiv E \\
& 1.1+\mathbb{Z}=\{\ldots, 0.1,1.1,2.1,3.1, \ldots\}=0.1+\mathbb{Z}=\equiv E+0.1
\end{aligned}
$$

This cycle repeats itself every $x=\ldots,-1,0,1,2, \ldots$ so that

$$
\begin{gathered}
\cdots=-1+\mathbb{Z}=0+\mathbb{Z}=1+\mathbb{Z}=2+\mathbb{Z}=\cdots \equiv E \\
\cdots=-0.9+\mathbb{Z}=0.1+\mathbb{Z}=1.1+\mathbb{Z}=2.1+\mathbb{Z}=\cdots \equiv E+0.1
\end{gathered}
$$

and so on. We can therefore "collapse" the set of all cosets $\{x+\mathbb{Z}\}, x \in \mathbb{R}$ by rewriting them as

$$
\mathbb{R} / \mathbb{Z}=\{E+\varphi\} \quad \varphi \in[0,1)
$$



Figure 1.9: Comparing the groups $\mathbb{R}$ and $\mathbb{R} / \mathbb{Z}$
Since $\mathbb{Z} \triangleleft \mathbb{R}$, we know that the cosets $E+\varphi$ form a group. To calculate the group multiplication $\star$, we notice that

$$
(E+0.9) \star(E+0.2)=(0.9+0.2) \mathbb{Z}=0.1 \mathbb{Z}=E+0.1
$$

That is, the group multiplication $\star$ is simply addition modulo 1 , denoted by $+_{1}$. This kind of addition "wraps around" 1. A well-known example is addition modulo 24 -if we were to add six hours to 20 h ( 8 pm ), we end up with 2 o'clock, not 26 o'clock:

$$
20+{ }_{24} 6=2 .
$$

Therefore, the quotient group $\mathbb{R} / \mathbb{Z}$ has the structure of a group formed by the set $[0,1)$ and addition modulo 1 :

$$
\mathbb{R} / \mathbb{Z}=[0,1)
$$

If we consider the normal subgroup $2 \pi \mathbb{Z}=\{\ldots,-2 \pi, 0,2 \pi, 4 \pi, \ldots\}$, the above argument produces

$$
\mathbb{R} / 2 \pi \mathbb{Z}=[0,2 \pi),
$$

the group formed by the set $[0,2 \pi)$ and addition modulo $2 \pi$.
Keep in mind that we have arrived at this answer using sheer "brute force." Although this method helps in visualizing the structure of the factor group, the result is not always readily apparent. In the next section, we will discover a slightly different way of determining factor groups via the first isomorphism theorem.

### 1.3 Homomorphisms

## Definition

We define a homomorphism $f: \mathrm{G} \rightarrow \mathrm{H}$ between two groups G and H as a mapping with the simple, but important property

$$
f\left(g \cdot g^{\prime}\right)=f(g) \cdot f\left(g^{\prime}\right) \quad \forall g, g^{\prime} \in \mathrm{G}
$$

This property is what separates homomorphisms from ordinary mappings. Homomorphisms are particularly useful since they ensure that the group multiplication of $G$ is preserved in H . Homomorphisms also have the following properties:

Proposition 1.1 Let $f: \mathrm{G} \rightarrow \mathrm{H}$ be a homomorphism. We have

1. $f(\mathbb{1})=\mathbb{1}$
2. $f\left(g^{-1}\right)=f(g)^{-1} \quad \forall g \in \mathrm{G}$.

Proof: We first show 1. by noticing that

$$
f(g)=f(g \cdot \mathbb{1})=f(g) \cdot f(\mathbb{1}) .
$$

Multiplying both sides by $f(g)^{-1}$ gives us $\mathbb{1}=f(\mathbb{1})$. To show 2 . we start with

$$
\mathbb{1}=f(\mathbb{1})=f\left(g \cdot g^{-1}\right)=f(g) \cdot f\left(g^{-1}\right) .
$$

Multiplying both sides by $f(g)^{-1}$ gives $f(g)^{-1}=f\left(g^{-1}\right)$.

We are already familiar with an example of a homomorphism:

$$
f: \mathrm{C}_{4} \rightarrow \mathbb{C}^{*} \quad f(a):=\mathrm{i}
$$

where $\mathbb{C}^{*}=\mathbb{C} \backslash\{0\}$ with multiplication of numbers as group operation. This homomorphism connects the abstract cyclic group $C_{4}$ with concrete complex numbers. We can easily calculate the action of $f$ on the other group members, since $f$ is homomorphic:

$$
\begin{aligned}
& f\left(a^{2}\right)=f(a \cdot a)=f(a) \cdot f(a)=\mathrm{i} \cdot \mathrm{i}=-1 \\
& f\left(a^{3}\right)=f\left(a \cdot a^{2}\right)=f(a) \cdot f\left(a^{2}\right)=\mathrm{i} \cdot-1=-\mathrm{i} \\
& f(\mathbb{1})=f\left(a^{4}\right)=f\left(a \cdot a^{3}\right)=f(a) \cdot f\left(a^{3}\right)=\mathrm{i} \cdot-\mathrm{i}=1 .
\end{aligned}
$$

## Subgroups Associated with Homomorphisms

When given a homomorphism $f: \mathrm{G} \rightarrow \mathrm{H}$, we can identify two important subgroups. The first one is called the image of $f$, written as $\operatorname{im} f \subset \mathrm{H}$. It is the set of all $h \in \mathrm{H}$ which are mapped by $f$ :

$$
\operatorname{im} f:=\{h \in \mathrm{H} \mid h=f(g) \quad g \in \mathrm{G}\}
$$

The kernel of $f$, written as $\operatorname{ker} f \subset \mathrm{G}$, is the set of all $g$ that are mapped into the identity element $\mathbb{1}$ of H :

$$
\operatorname{ker} f:=\left\{g \in \mathrm{G} \mid f(g)=\mathbb{1}_{\mathrm{H}}\right\}
$$

$\operatorname{im} f$ is sometimes written as $f(\mathrm{G})$, and $\operatorname{ker} f$ is sometimes written as $f^{-1}\left(\mathbb{1}_{H}\right)$. The image of a mapping is also referred to as the range.

Proposition 1.2 The kernel of a homomorphism $f: \mathrm{G} \rightarrow \mathrm{H}$ is a normal subgroup of G :

$$
\operatorname{ker} f \triangleleft G .
$$

Proof: Let $k \in \operatorname{ker} f$ so that $f(k)=\mathbb{1}$. We thus have

$$
f\left(g k g^{-1}\right)=f(g) f(k) f\left(g^{-1}\right)=f(g) \mathbb{1} f(g)^{-1}=\mathbb{1} \quad \forall g \in \mathrm{G} .
$$

We thus have $g k g^{-1} \in \operatorname{ker} f$ for all $g \in \mathrm{G}$.

## Isomorphisms

A homomorphism $f: \mathrm{G} \rightarrow \mathrm{H}$ is called injective if every $h \in \mathrm{H}$ is mapped by at most one $g \in \mathrm{G}$. We see that $f: \mathrm{C}_{4} \rightarrow \mathbb{C}$ is indeed injective:

$$
\begin{gathered}
a \mapsto \mathrm{i} \\
a^{2} \mapsto-1 \\
a^{3} \mapsto-\mathrm{i} \\
a^{4} \mapsto 1 .
\end{gathered}
$$

Alternatively, we can simply focus on the fact that the only element that maps to 1 is $a^{4}=\mathbb{1}$, that is, the kernel of $f$ is $\{\mathbb{1}\}$. This fact is guaranteed by the following proposition:

Proposition 1.3 Let $f: \mathrm{G} \rightarrow \mathrm{H}$ be a homomorphism. It is injective if and only if $\operatorname{ker} f=\{\mathbb{1}\}$.

$$
f \text { injective } \quad \Leftrightarrow \quad \text { ker } f=\{\mathbb{1}\} \text {. }
$$

Proof: First, assume $f$ is injective. Let $g \in \operatorname{ker} f$ so that $f(g)=\mathbb{1}$. By Proposition 1.1, we have

$$
f(g)=\mathbb{1}=f(\mathbb{1})
$$

Since $f$ is injective, we have $g=\mathbb{1}$. The above holds for all $g \in \mathrm{G}$, producing $\operatorname{ker} f=\{\mathbb{1}\}$. Conversely, assume $\operatorname{ker} f=\{\mathbb{1}\}$. Let $g, \gamma \in \mathrm{G}$ so that $f(g)=f(\gamma)$. If we consider the element $g \gamma^{-1}$, we see that

$$
f\left(g \gamma^{-1}\right)=f(g) f\left(\gamma^{-1}\right)=f(g) f(\gamma)^{-1}=\mathbb{1} .
$$

This shows that $g \gamma^{-1} \in \operatorname{ker} f$. Since, by assumption, the kernel only consists of the identity element, we see that $g \gamma^{-1}=\mathbb{1}$. Multiplying both sides by $\gamma$ from the right shows us that

$$
f(g)=f(\gamma) \quad \Leftrightarrow \quad g=\gamma
$$

Therefore, $f$ is injective.

A homomorphism $f: \mathrm{G} \rightarrow \mathrm{H}$ is called surjective if every $h \in \mathrm{H}$ is mapped by at least one $g \in \mathrm{G}$. Our previous homomorphism is not surjective, but this can be easily remedied by a minor cosmetic fix: replace H with the image of $f$ :

$$
f: \mathrm{C}_{4} \rightarrow\{\mathrm{i},-1,-\mathrm{i}, 1\}
$$

(Generally speaking, given a mapping $f: D \rightarrow C$ on a domain $D$, we can trivially make $f$ surjective by setting the range $C$ equal to im $f$.)
A homomorphism $f: \mathrm{G} \rightarrow \mathrm{H}$ is called bijective if it is both injective and surjective. A bijective homomorphism is also known as an isomorphism. Two groups $G$ and H are called isomorphic if there exists an isomorphism $f$ between them. We write $\mathrm{G} \cong \mathrm{H}$ then. Isomorphic groups are essentially the same. They may differ in form, but they have the same structure. Group theory cannot distinguish between them. Because the homomorphism $f: \mathrm{C}_{4} \rightarrow\{\mathrm{i},-1,-\mathrm{i}, 1\}$ is bijective, we have

$$
\mathrm{C}_{4} \cong\{\mathrm{i},-1,-\mathrm{i}, 1\}
$$

(The notion of isomorphism is very powerful, but surjectivity is often a nuisance since it is trivial to make any simple mapping surjective. For instance, the homomorphism $f: \mathrm{C}_{4} \rightarrow \mathbb{C}$ is not surjective, and therefore not technically an isomorphism. Instead, we'd constantly have to rewrite $f$ to make it surjective, leading to overly-explicit and drawn out mapping declarations like $f: \mathrm{C}_{4} \rightarrow\{\mathrm{i},-1,-\mathrm{i}, 1\} \subset \mathbb{C}$. For this reason, we will later introduce the notion of a faithful homomorphism. This is just a synonym for injective, so that when we say, for instance, that $f: \mathrm{C}_{4} \rightarrow \mathbb{C}$ is faithful, we understand that $f$ is an isomorphism on its image, not on all of $\mathbb{C}$.)
A bijective mapping is also called invertible, since for every bijective mapping $f$ there exists a bijective inverse mapping $f^{-1}$ with $f \circ f^{-1}=f^{-1} \circ f=i d$.
An injective mapping is also referred to as one-to-one. A bijective mapping is also referred to as one-to-one correspondence. An isomorphism $f: \mathrm{G} \rightarrow \mathrm{G}$, which maps a group onto itself, is called an automorphism. The set of automorphisms of a group, denoted by Aut G, forms a group under composition.

## The First Isomorphism Theorem

We will now concern ourselves with the first isomorphism theorem for groups. Using this theorem, we will have another way of determining factor groups.

Theorem 1.4 (First Isomorphism Theorem) Let G and H be groups, and $f: \mathrm{G} \rightarrow \mathrm{H}$ be a homomorphism. We have

$$
\mathrm{G} / \operatorname{ker} f \cong \operatorname{im} f .
$$

Proof: See Jones, §2.4.
(There are two further isomorphism theorems, which are more intricate and less prominent than the first one given above. They are beyond the scope of this script and won't be included.)
Let's return to the example of $\mathbb{R} / \mathbb{Z}$. Consider the mapping $f: \mathbb{R} \rightarrow \mathbb{C}^{*}$ defined by

$$
f(x)=e^{i x}
$$

where $\mathbb{C}^{*}$ is the group formed by the set $\mathbb{C} \backslash\{0\}$ under scalar multiplication. We see that $f$ is a homomorphism because

$$
f(x+y)=e^{i(x+y)}=e^{i x} \cdot e^{i y}=f(x) \cdot f(y) .
$$

Its image and kernel are

$$
\begin{aligned}
\operatorname{im} f & =S^{1}=\{z \in \mathbb{C}| | z \mid=1\} \\
\operatorname{ker} f & =2 \pi \mathbb{Z}=\{\cdots-2 \pi, 0,2 \pi, 4 \pi, \ldots\}
\end{aligned}
$$

where $S^{1}$ is the unit circle. The first isomorphism theorem gives us

$$
\mathbb{R} / 2 \pi \mathbb{Z} \cong S^{1}
$$

Recall in the last section where we showed by "brute force" that

$$
\mathbb{R} / 2 \pi \mathbb{Z} \cong[0,2 \pi) .
$$

A little thought will convince the reader that $S^{1}$ with multiplication as group operation and $[0,2 \pi)$ with addition modulo 2 as group operation are indeed isomorphic. (Consider the mapping $\varphi \mapsto e^{i \varphi}$ and check that it is an isomorphism.)
The first isomorphism theorem provides a way to construct factor groups. However, do realize that we need an explicit homomorphism for that, and arriving at one isn't always a straightforward task. Some texts might even grant the reader with a conveniently given homomorphism, providing one with the impression that determining factor groups is incredibly easy. Notice that the previous example $\mathbb{R} / \mathbb{Z}=S^{1}$ is guilty of this. If a similar example is difficult to grasp intuitively, try to write down the cosets to understand the homomorphism as well as the structure of the factor group.

### 1.4 Algebras

Aside from groups, we will be dealing with another kind of important mathematical structure in this text, namely that of an algebra.

## Definition

We will begin with a mathematical structure which should already be familiar to the reader. A vector space

$$
(V,+, \cdot)
$$

over a field $\mathbb{K}$ is a set $V$ together with two operations vector addition + and scalar multiplication • which satisfy the following properties: Vector addition is associative, commutative, has an identity $\mathbf{0}$, and is invertible. Scalar multiplication is associative, has an identity 1 , is distributive over vector addition, and is distributive over field addition. An algebra

$$
A=(V,+, \cdot, \times)
$$

is simply a vector space $V$ over a field $\mathbb{K}$, with an extra operation $\times$ called algebra multiplication which is bilinear:

$$
\begin{array}{rl}
(\mathbf{x}+\mathbf{y}) \times \mathbf{z}=\mathbf{x} \times \mathbf{z}+\mathbf{y} \times \mathbf{z} & \mathbf{x} \times(\mathbf{y}+\mathbf{z})=\mathbf{x} \times \mathbf{y}+\mathbf{x} \times \mathbf{z} \\
(a \mathbf{x}) \times \mathbf{y}=a(\mathbf{x} \times \mathbf{y}) & \mathbf{x} \times(b \mathbf{y})=b(\mathbf{x} \times \mathbf{y})
\end{array}
$$

Notice that we don't require the multiplication $\times$ to be commutative or even associative. In general,

$$
\mathrm{x} \times \mathrm{y} \neq \mathrm{y} \times \mathrm{x} \quad \mathrm{x} \times(\mathrm{y} \times \mathrm{z}) \neq(\mathrm{x} \times \mathrm{y}) \times \mathrm{z}
$$

In fact, if a multiplication isn't associative, we usually (but not always) denote it with some sort of brackets like [, ] instead of a lone symbol like $\times$ :

$$
[\mathbf{x}, \mathbf{y}] \neq[\mathbf{y}, \mathbf{x}] \quad[\mathbf{x},[\mathbf{y}, \mathbf{z}]] \neq[[\mathbf{x}, \mathbf{y}] \mathbf{z}]
$$

Therefore, we introduce the notion of an associative algebra for the special case of an algebra formed under an associative algebra multiplication.

## Structure Constants

Say we have a basis $\left\{\mathbf{e}_{1}, \ldots, \mathbf{e}_{n}\right\}$ of the algebra $A$, where $\mathbf{x}, \mathbf{y} \in A$ are given by linear combinations of the basis vectors:

$$
\mathbf{x}=x^{i} \mathbf{e}_{i} \quad \mathbf{y}=y^{j} \mathbf{e}_{j} \quad x^{i}, y^{j} \in \mathbb{K}
$$

We see that the product of any two vectors is given by

$$
\mathbf{x} \times \mathbf{y}=\left(x^{i} \mathbf{e}_{i}\right) \times\left(y^{j} \mathbf{e}_{j}\right)=x^{i} y^{i}\left(\mathbf{e}_{i} \times \mathbf{e}_{j}\right) .
$$

Therefore, if we can describe how the algebra multiplication acts on the basis

$$
\mathbf{e}_{i} \times \mathbf{e}_{j},
$$

then we have essentially described how the algebra multiplication acts on all vectors $\mathbf{x}, \mathbf{y} \in A$.
Because algebra multiplication is closed, we know that $\mathbf{e}_{i} \times \mathbf{e}_{j} \in A$. That is, it must be a linear combination of the basis vectors $\left\{\mathbf{e}_{i}, \ldots, \mathbf{e}_{n}\right\}$ :

$$
\mathbf{e}_{i} \times \mathbf{e}_{j}=c_{i j}{ }^{k} \mathbf{e}_{k}
$$

The numbers $c_{i j}^{k} \in \mathbb{K}$ are called structure constants. Knowing these constants may allow us to identify an algebra up to isomorphism, but they depend on the basis chosen.

## Lie Algebras

One important example of an algebra is a Lie algebra. Its algebra multiplication is denoted by [, ] and is referred to as the Lie bracket. It satisfies the following properties:

- Bilinearity: $[a \mathbf{x}+b \mathbf{y}, \mathbf{z}]=a[\mathbf{x}, \mathbf{z}]+b[\mathbf{y}, \mathbf{z}]$
- Antisymmetry: $[\mathbf{x}, \mathbf{y}]=-[\mathbf{y}, \mathbf{x}]$
- The Jacobi Identity: $[\mathbf{x},[\mathbf{y}, \mathbf{z}]]+[\mathbf{y},[\mathbf{z}, \mathbf{x}]]+[\mathbf{z},[\mathbf{x}, \mathbf{y}]]=0$

Notice that the Lie bracket is neither commutative nor associative.
For example, consider the vector space $M$ of $n \times n$ matrices. Notice that there are two possible algebra multiplications. First, we have matrix multiplication $\cdot$, which is associative:

$$
\left(\mathbf{M}_{1} \cdot \mathbf{M}_{2}\right) \cdot \mathbf{M}_{3}=\mathbf{M}_{1} \cdot\left(\mathbf{M}_{2} \cdot \mathbf{M}_{3}\right) .
$$

There's also the commutator

$$
\left[\mathbf{M}_{1}, \mathbf{M}_{2}\right]:=\mathbf{M}_{1} \cdot \mathbf{M}_{2}-\mathbf{M}_{2} \cdot \mathbf{M}_{1}
$$

which is not associative but does satisfy the defining properties of a Lie bracket. Therefore, the vector space of matrices can form two different algebras: the associative algebra ( $M, \cdot$ ) and the Lie algebra ( $M,[$,$] ).$
Another example of an algebra is the vector space $\mathbb{R}^{3}$ with respect to the canonical basis $\left\{\mathbf{e}_{1}, \mathbf{e}_{2}, \mathbf{e}_{3}\right\}$ and the Euclidean cross product $\times$ as algebra multiplication:

$$
A=\left(\mathbb{R}^{3}, \times\right)
$$

The cross product also satisfies the three properties of a Lie bracket, making $A$ a Lie algebra. We already know the structure constants of this algebra, since

$$
\mathbf{e}_{i} \times \mathbf{e}_{j}=\epsilon_{i j}^{k} \mathbf{e}_{k}
$$

That is, the structure constants are given by the totally antisymmetric tensor of rank 3

$$
c_{i j}^{k}=\epsilon_{i j}^{k}=\left\{\begin{aligned}
1 & \text { for } \operatorname{sgn}(\sigma)=1 \\
-1 & \text { for } \operatorname{sgn}(\sigma)=-1 \\
0 & \text { otherwise }
\end{aligned}\right.
$$

where $\sigma(\{1,2,3\})=\{i, j, k\}$.

## Homomorphisms Between Algebras

A homomorphism $f: A \rightarrow B$ can be defined on two algebras $A$ and $B$ as a mapping with the property

$$
f[X, Y]=[f X, f Y] \quad \forall X, Y \in A
$$

The image im $f$ of a homomorphism $f: A \rightarrow B$ is defined as

$$
\operatorname{im} f:=\{b \in B \mid b=f(a) \quad a \in A\} .
$$

The kernel ker $f$ of a homomorphism $f: A \rightarrow B$ is defined as

$$
\operatorname{ker} f:=\{a \in A \mid f(a)=\mathbf{0}\}
$$

A homomorphism $f: A \rightarrow B$ is called injective if every $b \in B$ is mapped by at most one $a \in A$. It is called surjective if every $b \in B$ is mapped by at least one $a \in A$. A homomorphism $f: A \rightarrow B$ is called bijective if it is both injective and surjective.
There holds an isomorphism theorem for algebras similar to that for groups:
Theorem 1.5 (First Isomorphism Theorem for Algebras) Let $A$ and $B$ be algebras, and $f: A \rightarrow B$ be a homomorphism. We have

$$
A / \operatorname{ker} f \cong \operatorname{im} f
$$

## Chapter 2

## Representations

We have already seen that the group $C_{4}$ can be mapped into the complex plane $\mathbb{C}$ by the homomorphism

$$
f(a)=\mathrm{i} .
$$

That is, we can associate group elements with points on a complex space. However, we need something more powerful than $f$ if we want to take advantage of the symmetries that groups portray. This is where representations come into play.

### 2.1 General Representations

## The General Linear Group $\mathrm{GL}(V)$

Let $V$ denote a vector space over the field $\mathbb{K}$, where $\mathbb{K}=\mathbb{R}, \mathbb{C}$. We define the general linear group $\mathrm{GL}(V)$ as the set of all bijective linear transformations of $V$ :

$$
\mathrm{GL}(V):=\{T: V \rightarrow V \mid T \text { bijective linear }\} .
$$

(Because the bijective linear transformations $T: V \rightarrow V$ are just the automorphisms of $V$, we could also write $\operatorname{Aut}(V)$ instead of $\mathrm{GL}(V)$.)
This space forms a group under composition $\circ . \mathrm{GL}(V)$ contains several subgroups. The most prominent ones are the so-called classical subgroups:
Let $s: V \times V \rightarrow \mathbb{K}$ be a nondegenerate symmetric bilinear form: $s(u, v)=s(v, u)$. The orthogonal group $\mathrm{O}(V) \subset \mathrm{GL}(V)$ consists of the transformations that preserve $s$ :

$$
\mathrm{O}(V, s):=\{T \in \mathrm{GL}(V) \mid s(T u, T v)=s(u, v)\} .
$$

In case $\mathbb{K}=\mathbb{C}$, let $h: V \times V \rightarrow \mathbb{C}$ be a nondegenerate hermitean bilinear form so that $h(u, v)=h(v, u)^{*}$. The unitary group $\mathrm{U}(V) \subset \mathrm{GL}(V)$ consists of the transformations that preserve $h$ :

$$
\mathrm{U}(V, h):=\{T \in \mathrm{GL}(V) \mid h(T u, T v)=h(u, v)\} .
$$

Let $a: V \times V \rightarrow \mathbb{K}$ be a nondegenerate anti-symmetric bilinear form so that $a(u, v)=$ $-a(v, u)$. The symplectic group $\mathrm{Sp}(V) \subset \mathrm{GL}(V)$ consists of the transformations that preserve $a$ :

$$
\operatorname{Sp}(V, a):=\{T \in \mathrm{GL}(V) \mid a(T u, T v)=a(u, v)\}
$$

These admittedly abstract definitions will have more significance as we consider matrix representations in the next subsection.

## Representations

A representation is a homomorphism $\mathrm{D}: \mathrm{G} \rightarrow \mathrm{GL}(V)$ that maps a group element $g \in \mathrm{G}$ to a linear transformation $\mathrm{D}(g): V \rightarrow V$. Because representations are homomorphisms, don't forget that

$$
\mathrm{D}\left(g \cdot g^{\prime}\right)=\mathrm{D}(g) \cdot \mathrm{D}\left(g^{\prime}\right)
$$

So why are representations so important? Well, recall the homomorphism

$$
f: C_{4} \rightarrow \mathbb{C}^{*}, \quad f(a)=\mathrm{i}
$$

This mapping is not a representation since it maps into $\mathbb{C}^{*}$, not $\mathrm{GL}(V)$. That is, it simply assigns each group element to a point in the complex plane with the origin removed. Now consider the representation $\mathrm{D}: C_{4} \rightarrow \mathrm{GL}(\mathbb{C})$

$$
\mathrm{D}(a)=\mathrm{i}
$$

$D(a)$ may be some point in the complex plane, but it is better viewed as a linear transformation $\mathrm{D}(a): \mathbb{C} \rightarrow \mathbb{C}$, which rotates any $z \in \mathbb{C}$ by $\frac{\pi}{2}$. In this way, the concept of rotation provided by the abstract group element $a$ is "represented" in the complex plane by D

$$
\mathrm{D}(a) z=\mathrm{i} \cdot z
$$



Figure 2.1: The linear transformation $\mathrm{D}(a)$, which rotates $z$
In fact, we can extend this representation to the dihedral group of order four. Consider now the representation $\mathrm{D}: \mathrm{D}_{4} \rightarrow \mathbb{C}$ defined by

$$
\mathrm{D}(a) z=\mathrm{i} z \quad \mathrm{D}(b) z=z^{*} .
$$



Figure 2.2: The linear transformation $\mathrm{D}(b)$, which reflects $z$

We see that the concept of reflection provided by the abstract group element $b$ is also "represented" in the complex plane by D. Notice that $\mathrm{D}(b)$ cannot correspond to any complex number; it can only be viewed as a transformation.
By mapping into linear transformations, representations give groups the power to manipulate and transform any kind of vector space.

## Terminology

A representation D of a group $G$ is called faithful if it is injective, that is, if it is an isomorphism on its image. In this case, group elements $g$ and transformations $\mathrm{D}(g)$ are indistinguishable, and we essentially have an identical copy of $G$ as a subgroup of $\mathrm{GL}(V)$. The representation of the group $\mathrm{D}_{4}$ given earlier is indeed faithful. Therefore, the transformations $\mathrm{D}(a)$ and $\mathrm{D}(b)$ are essentially the same as the group elements $a$ and $b$. The carrier space refers to the vector space $V$ whose linear transformations are mapped by $D$. The representation of $D_{4}$ given earlier has a carrier space $\mathbb{C}$. Many different carrier spaces are possible. For instance, the space of quantum-mechanical wavefunctions $L^{2}\left(\mathbb{R}^{3}, \mathrm{~d} \mu\right)$ is a common carrier space in quantum mechanics. (In such a case, the linear transformations produced are commonly known as operators.)
A representation D is said to have a dimension equal to that of its carrier space $V$,

$$
\operatorname{dim}(\mathrm{D})=\operatorname{dim}(V)
$$

A representation is called real or complex, depending on whether its carrier space is real or complex.
If $V$ carries a scalar product $\langle\cdot, \cdot\rangle$ on $V$, then we say that a representation D is orthogonal (real case) or unitary (complex case) if

$$
\langle\mathrm{D}(g) u, \mathrm{D}(g) v\rangle=\langle u, v\rangle,
$$

or equivalently,

$$
\langle\mathrm{D}(g) u, v\rangle=\left\langle u, \mathrm{D}(g)^{-1} v\right\rangle .
$$

## Invariance and Reducibility

Recall that the group $C_{4}$ corresponds to the symmetry of a square. That is, a rotation of $90^{\circ}$ leaves a square unchanged. In other words, the square is left invariant under a $\mathrm{C}_{4}$ transformation.
Let $\mathrm{T}: \mathrm{G} \rightarrow \mathrm{GL}(V)$ be a representation. A subspace $S \subset V$ is called invariant under T if

$$
s \in S \quad \Rightarrow \quad T(g) s \in S \quad \forall g \in G
$$

That is, every transformation $\mathrm{T}\left(g_{1}\right), \mathrm{T}\left(g_{2}\right), \ldots$ promises not to kick any vector $s$ out of its subspace $S$.
A representation $\mathrm{T}: \mathrm{G} \rightarrow \mathrm{GL}(V)$ is called reducible if there exists a non-trivial subspace $S$ (i.e. $S \neq\{0\}$ and $S \neq V$ ) that is invariant under T.
A representation $\mathrm{T}: G \rightarrow \mathrm{GL}(V)$ is called completely reducible if every invariant subspace $S$ has an invariant complement $S^{\prime}$.
As their names imply, reducible representations can still be "reduced" into components. Let T be a completely reducible representation with an invariant subspace $S \subset V$. Just as $V$ consists of a direct sum of a subspace and its orthogonal complement,

$$
V=S \oplus S^{\prime}
$$

a completely reducible representation T takes on the form

$$
\mathrm{T}=\mathrm{T}_{S} \oplus \mathrm{~T}_{S^{\perp}}
$$

Perhaps $\mathrm{T}_{S}$ can be completely reduced even further by invariant subspaces $R \subset S$, or perhaps there aren't any subspaces of $S$ (or of $S^{\perp}$ ) that are invariant under T. In this case, we would call $\mathrm{T}_{S}$ (or $\mathrm{T}_{S^{\perp}}$ ) irreducible:
A representation $\mathrm{T}: \mathrm{G} \rightarrow \mathrm{GL}(V)$ is called irreducible if it is not reducible. That is, the only invariant subspaces $S \subseteq V$ are trivial.

### 2.2 Matrix Representations

Another kind of representation is so useful, we will devote an entire section to it. These use the carrier spaces $\mathbb{R}^{n}$ or $\mathbb{C}^{n}$. Recall from linear algebra that any linear transformation over these spaces can be expressed as a matrix $\mathbf{M}$ by transforming the basis $\mathbf{e}_{i}$ :

$$
\mathbf{M}_{j}^{i} \mathbf{e}_{i}=\hat{\mathrm{T}} \mathbf{e}_{j}
$$

Thus, any operator $\mathrm{T}(g)$ can be identified as an $n \times n$ matrix defined by

$$
\mathrm{T}(g) \mathbf{e}_{j}=\mathrm{D}(g)^{i}{ }_{j} \mathbf{e}_{i} .
$$

We call this representation D a matrix representation. These kinds of representations allow us to tackle group theory using the reliable techniques learned in linear algebra.

## The General Linear Group and its Subgroups in Matrix Form

We define the general linear groups of invertible (i.e. non-zero determinant) matrices:

$$
\begin{aligned}
& \mathrm{GL}(n, \mathbb{R}):=\left\{M \in \mathbb{R}^{n \times n} \mid \operatorname{det} M \neq 0\right\} \\
& \mathrm{GL}(n, \mathbb{C}):=\left\{M \in \mathbb{C}^{n \times n} \mid \operatorname{det} M \neq 0\right\} .
\end{aligned}
$$

Notice that these matrix groups correspond to the groups $\mathrm{GL}\left(\mathbb{R}^{n}\right)$ and $\mathrm{GL}\left(\mathbb{C}^{n}\right)$ of linear transformations. In particular, we define the special linear groups of matrices with unit determinant:

$$
\begin{aligned}
& \mathrm{SL}(n, \mathbb{R}):=\left\{M \in \mathbb{R}^{n \times n} \mid \operatorname{det} M=1\right\} \\
& \mathrm{SL}(n, \mathbb{C}):=\left\{M \in \mathbb{C}^{n \times n} \mid \operatorname{det} M=1\right\} .
\end{aligned}
$$

The Euclidean dot product • acts as a symmetric bilinear form over $\mathbb{R}^{n}$. We see that it is preserved by orthogonal matrices:

$$
M \vec{x} \cdot M \vec{y}=\vec{x} \cdot \vec{y} \quad \Leftrightarrow \quad M^{\mathrm{T}} M=\mathbb{1} .
$$

Thus, we define the group of orthogonal matrices and its special subgroup:

$$
\begin{aligned}
\mathrm{O}(n) & :=\left\{M \in \mathrm{GL}(n, \mathbb{R}) \mid M^{\mathrm{T}} M=\mathbb{1}\right\} \\
\mathrm{SO}(n) & :=\{M \in \mathrm{O}(n) \mid \operatorname{det} M=1\} .
\end{aligned}
$$

Notice that the matrix group $\mathrm{O}(n)$ corresponds to the group $\mathrm{O}\left(\mathbb{R}^{n}, \cdot\right)$ of linear transformations, and that an orthogonal matrix $\mathbf{M} \in O(n)$ preserves the length of a vector $\vec{x} \in \mathbb{R}^{n}$

$$
\|\mathbf{M} \vec{x}\|^{2}=\mathbf{M} \vec{x} \cdot \mathbf{M} \vec{x}=\vec{x} \cdot \vec{x}=\|\vec{x}\|^{2}
$$

The hermitean scalar product $\langle$,$\rangle acts as a hermitean bilinear form over \mathbb{C}^{n}$. We see that it is preserved by unitary matrices:

$$
\langle\mathbf{M} u, \mathbf{M} v\rangle=\langle u, v\rangle \quad \Leftrightarrow \quad \mathbf{M}^{\dagger} \mathbf{M}=\mathbb{1} .
$$

Thus, we define the group of unitary matrices and its special subgroup as

$$
\begin{aligned}
\mathrm{U}(n) & :=\left\{\mathbf{M} \in \mathrm{GL}(n, \mathbb{C}) \mid \mathbf{M}^{\dagger} \mathbf{M}=\mathbb{1}\right\} \\
\mathrm{SU}(n) & :=\{\mathbf{M} \in \mathbf{U}(n) \mid \operatorname{det} \mathbf{M}=1\}
\end{aligned}
$$

Notice that the matrix group $\mathrm{U}(n)$ corresponds to the group $\mathrm{U}\left(\mathbb{C}^{n},\langle\rangle,\right)$ of linear transformations, and that a unitary matrix $\mathbf{M} \in U(n)$ preserves the norm of a complex number $v \in \mathbb{C}^{n}$

$$
\|\mathbf{M} v\|^{2}=\langle\mathbf{M} v, \mathbf{M} v\rangle=\langle v, v\rangle=\|v\|^{2}
$$

## Terminology

A homomorphism $\mathrm{D}: \mathrm{G} \rightarrow \mathrm{GL}(n, \mathbb{R})$ is called a real matrix representation and a homomorphism $\mathrm{D}: \mathrm{G} \rightarrow \mathrm{GL}(n, \mathbb{C})$ is called a complex matrix representation. They allow us to qualitatively express a group's action on $\mathbb{R}^{n}$ or $\mathbb{C}^{n}$ using matrices.
The number $n$ refers to the dimension of a representation, which is equal to the dimension of the carrier space.

## Characters and Equivalent Representations

Although matrices allow us to concretely express a linear transformation on $\mathbb{R}^{n}$ or $\mathbb{C}^{n}$, they do have a certain weakness: they depend on the basis chosen. A single transformation corresponds to a set of infinitely many matrices which are related by a similarity transformation

$$
\begin{array}{lll}
\mathbf{M}^{\prime}=\mathbf{S M S}^{-1} & \mathbf{S} \in \mathbf{S O}(n) & \text { (real case) } \\
\mathbf{M}^{\prime}=\mathbf{S M S}^{-1} & \mathbf{S} \in \mathrm{SU}(n) & \text { (complex case) }
\end{array}
$$

where the matrix $\mathbf{S}$ transforms an oriented, orthonormal basis. (For an arbitary basis, $\mathbf{S} \in \mathrm{GL}(n, \mathbb{R}), \mathrm{GL}(n, \mathbb{C})$.) These similar matrices may be infinite in number, but they are all related by a simple fact - their traces are the same:

$$
\operatorname{tr} \mathbf{M}^{\prime}=\operatorname{tr} \mathbf{S M S}^{-1}=\operatorname{tr} \mathbf{M}
$$

Thus, we can define an equivalence relation between similar matrices:

$$
\mathbf{M}^{\prime} \sim \mathbf{M} \quad: \Leftrightarrow \quad \operatorname{tr} \mathbf{M}^{\prime}=\operatorname{tr} \mathbf{M}
$$

We use the same idea for matrix representations. Let D be a matrix representation of a group $G$. The mapping $\chi: G \rightarrow \mathbb{C}$ defined by

$$
\chi(g):=\operatorname{tr} \mathrm{D}(g)
$$

is called the character of the matrix representation D .
(Some authors additionally define $\chi$ to be a set of traces of all matrices $\mathrm{D}(g)$ :

$$
\chi:=\{\chi(g)\}_{g \in \mathrm{G}}:=\{\operatorname{tr} \mathrm{D}(g)\}_{g \in \mathrm{G}}
$$

where $\chi(g)=\operatorname{tr} \mathrm{D}(g)$.)
Just as we have established an equivalence relation among matrices, we can establish an equivalence relation among matrix representations in the same way. Two matrix representations $\mathrm{D}, \mathrm{D}^{\prime}: \mathrm{G} \rightarrow \mathrm{GL}(n, \mathbb{K})$ are called equivalent if their characters are identical:

$$
\mathrm{D} \sim \mathrm{D}^{\prime} \quad: \Leftrightarrow \quad \chi_{\mathrm{D}}=\chi_{\mathrm{D}^{\prime}}
$$

Proposition 2.1 Two n-dimensional matrix representations D and $\mathrm{D}^{\prime}$ are equivalent if and only if

$$
\mathrm{D}^{\prime}(g)=\mathbf{S D}(g) \mathbf{S}^{-1} \quad \forall g \in \mathrm{G}
$$

for some n-dimensional invertible matrix $\mathbf{S}$.

## Reducibility

The abstract definition given earlier of a reducible representation $T: G \rightarrow G L(V)$ can be expressed more concretely if we consider a corresponding matrix representation $\mathrm{D}: \mathrm{G} \rightarrow$ $\mathrm{GL}(n, \mathbb{K})$.

Proposition 2.2 Any reducible matrix representation is equivalent to a matrix representation D that is of the form

$$
\mathrm{D}(g)=\left[\begin{array}{cc}
\mathrm{D}_{1}(g) & \mathrm{E}(g) \\
0 & \mathrm{D}_{2}(g)
\end{array}\right]
$$

while any completely reducible matrix representation is equivalent to a matrix representation D of the form

$$
\mathrm{D}(g)=\left[\begin{array}{cc}
\mathrm{D}_{1}(g) & 0 \\
0 & \mathrm{D}_{2}(g)
\end{array}\right] .
$$

Proof: Let $S$ be a proper invariant subspace and let $S^{\prime}$ be a complement in $V$, i.e., $V=S \oplus S^{\prime}$. We write

$$
\mathbf{v}=\left[\begin{array}{l}
s \\
t
\end{array}\right], \quad s \in S, t \in S^{\prime}
$$

Since D is reducible, we have

$$
S \ni \mathbf{s}=\left[\begin{array}{l}
s \\
0
\end{array}\right] \quad \Rightarrow \quad \mathrm{D}(g) \mathbf{s}=\left[\begin{array}{ll}
\mathrm{D}(g)^{1}{ }_{1} & \mathrm{D}(g)^{1}{ }_{2} \\
\mathrm{D}(g)^{2}{ }_{1} & \mathrm{D}(g)^{2}{ }_{2}
\end{array}\right]\left[\begin{array}{l}
s \\
0
\end{array}\right]=\left[\begin{array}{lll}
\mathrm{D}(g)^{1}{ }_{1} & s \\
\mathrm{D}(g)^{2}{ }_{1} & s
\end{array}\right] \in S .
$$

Since $\mathrm{D}(g) s$ cannot have an $S^{\prime}$-component, $\mathrm{D}(g)^{2}{ }_{1}=0$. If D is completely reducible, we can choose $S^{\prime}$ to be invariant, too. Then the same argument holds for $S^{\prime}$, hence $\mathrm{D}(\mathrm{g})^{1}{ }_{2}=0$.

Recall from linear algebra that a vector space $V$ can be decomposed into a direct sum of two subspaces

$$
V=S_{1} \oplus S_{2} \quad: \Leftrightarrow \quad V \ni v=\left[\begin{array}{l}
s_{1} \\
s_{2}
\end{array}\right]
$$

provided that $S_{1} \cap S_{2}=0$ and $\operatorname{dim} S_{1}+\operatorname{dim} S_{2}=\operatorname{dim} V$. Similarly, we can express a completely reducible representation by

$$
\mathrm{D}=\mathrm{D}_{1} \oplus \mathrm{D}_{2} \quad: \Leftrightarrow \quad \mathrm{D}(g)=\left[\begin{array}{cc}
\mathrm{D}_{1}(g) & 0 \\
0 & \mathrm{D}_{2}(g)
\end{array}\right] \quad \forall g \in G
$$

$\mathrm{D}_{1}$ and $\mathrm{D}_{2}$ are also representations, where $\mathrm{D}_{1}(g): S_{1} \rightarrow S_{1}$ and $\mathrm{D}_{2}: S_{2} \rightarrow S_{2}$.
If these two representations are again completely reducible, we can continue reducing until we end up with a sum of irreducible representations

$$
\mathrm{D}=\mathrm{D}_{1} \oplus \mathrm{D}_{2} \oplus \cdots \oplus \mathrm{D}_{N}=\bigoplus_{\nu=1}^{N} \mathrm{D}_{\nu}
$$

### 2.3 Irreducible Representations

In this section, we will explore some properties of irreducible representations.
Proposition 2.3 Every reducible unitary representation T is completely reducible.
Proof: Let T: $\mathrm{G} \rightarrow \mathrm{GL}(V)$ be reducible and unitary, and let $S \subset V$ be a nontrivial subspace invariant under T . We show that $S^{\perp}$ is invariant under T .

$$
\begin{array}{rll}
s \in S, t \in S^{\perp} & \Rightarrow & 0=\langle s, t\rangle \\
S \text { invariant } & \Rightarrow & 0=\langle\mathrm{T}(g) s, t\rangle \\
\mathrm{T} \text { unitary } & \Rightarrow & 0=\left\langle s, \mathrm{~T}^{-1}(g) t\right\rangle=\left\langle s, \mathrm{~T}\left(g^{-1}\right) t\right\rangle \\
& \Rightarrow & \mathrm{T}\left(g^{-1}\right) t \in S^{\perp}
\end{array}
$$

Writing $h \equiv g^{-1}$, we see that

$$
t \in S^{\perp} \quad \Rightarrow \quad \mathrm{T}(h) t \in S^{\perp} \quad \forall h \in \mathrm{G}
$$

Thus, $S^{\perp}$ is invariant under T , and we can iterate the argument to conclude that T is completely reducible.

Theorem 2.4 (Maschke's Theorem) All reducible representations of a finite group are completely reducible.

Proof: We define a scalar product on $V$ by

$$
\{v, w\}:=\frac{1}{\text { ord } \mathrm{G}} \sum_{g \in \mathrm{G}}\langle\mathrm{~T}(g) v, \mathrm{~T}(g) w\rangle
$$

Applying this shows

$$
\begin{aligned}
\left\{\mathrm{T}\left(g^{\prime}\right) v, \mathrm{~T}\left(g^{\prime}\right) w\right\} & =\frac{1}{\text { ord } \mathrm{G}} \sum_{g \in \mathrm{G}}\left\langle\mathrm{~T}(g) \mathrm{T}\left(g^{\prime}\right) v, \mathrm{~T}(g) \mathrm{T}\left(g^{\prime}\right) w\right\rangle \\
& =\frac{1}{\operatorname{ord} \mathrm{G}} \sum_{g \in \mathrm{G}}\left\langle\mathrm{~T}\left(g g^{\prime}\right) v, \mathrm{~T}\left(g g^{\prime}\right) w\right\rangle \\
& =\frac{1}{\text { ord } \mathrm{G}} \sum_{h \in \mathrm{G}}\langle\mathrm{~T}(h) v, \mathrm{~T}(h) w\rangle \\
\left\{\mathrm{T}\left(g^{\prime}\right) v, \mathrm{~T}\left(g^{\prime}\right) w\right\} & =\{v, w\},
\end{aligned}
$$

where $g^{\prime} \in \mathrm{G}$ and $g g^{\prime}=h \in \mathrm{G}$, and $\sum_{g \in \mathrm{G}}=\sum_{h \in \mathrm{G}}$. We see that T is unitary with respect to the scalar product defined above. According to Proposition 2.3, T is thus completely reducible.

Theorem 2.5 (Schur's First Lemma) Let $\mathrm{T}: \mathrm{G} \rightarrow \mathrm{GL}(V)$ be a complex irreducible representation, and let $\hat{\mathrm{B}}: V \rightarrow V$ be some linear operator that commutes with all $\mathrm{T}(\mathrm{g})$ :

$$
\hat{\mathrm{B}} \mathrm{~T}(g)=\mathrm{T}(g) \hat{\mathrm{B}} \quad \forall g \in \mathrm{G}
$$

Then $\hat{\mathrm{B}}=\lambda \mathbb{1}$ for $\lambda \in \mathbb{C}$.
Proof: $\hat{\mathrm{B}}$ possesses an eigenvector $b$. Let $\lambda$ denote the corresponding eigenvalue,

$$
\hat{\mathrm{B}} b=\lambda b
$$

Then

$$
\hat{\mathrm{B}} \cdot \mathrm{~T}(g) b=\mathrm{T}(g) \hat{\mathrm{B}} b=\lambda \cdot \mathrm{T}(g) b \quad \forall g \in \mathrm{G} .
$$

That is, $\mathrm{T}(g) b$ is also an eigenvector of $\hat{\mathrm{B}}$ with eigenvalue $\lambda$. Hence, the subspace $E_{\lambda}$ of all eigenvectors of $\hat{\mathrm{B}}$ with eigenvalue $\lambda$ is invariant under T . Since T is irreducible, $E_{\lambda}$ must be one of the trivial subspaces $V$ or $\{0\}$. Since eigenvectors are by definition nonzero, we rule out $\{0\}$ and conclude that $E_{\lambda}=V$. That is, all vectors $v \in V$ are eigenvectors to $\hat{\mathrm{B}}$ with eigenvalue $\lambda$ :

$$
\hat{\mathrm{B}} v=\lambda v \quad \forall v \in V
$$

Thus, $\hat{\mathrm{B}}=\lambda \mathbb{1}$.

The assumption that $V$ be complex is crucial; Schur's first lemma does not hold for real representations. A counterexample is provided by the identical representation of the rotation group $\mathrm{SO}(2)$ on $V=\mathbb{R}^{2}$, see $\S 3.1$. While all group elements commute with one another, the representation is obviously irreducible. The reader is encouraged to answer the following question. Identifying $\mathrm{SO}(2)$ with the unitary group $\mathrm{U}(1)$ and the representation space $\mathbb{R}^{2}$ with $\mathbb{C}$, this representation gets identified with a complex irreducible representation of $\mathbf{U}(1)$. Why does Schur's first lemma hold now?

Theorem 2.6 (Schur's Second Lemma) Let $\mathrm{T}_{1}: \mathrm{G} \rightarrow \mathrm{GL}\left(V_{1}\right)$ and $\mathrm{T}_{2}: \mathrm{G} \rightarrow \mathrm{GL}\left(V_{2}\right)$ be two irreducible representations of G , and let $\hat{\mathrm{B}}: V_{1} \rightarrow V_{2}$ be a linear operator satisfying

$$
\hat{\mathrm{B}} \mathrm{~T}_{1}(g)=\mathrm{T}_{2}(g) \hat{\mathrm{B}} \quad \forall g \in \mathrm{G} .
$$

Then $\hat{\mathrm{B}}$ is an isomorphism, or $\hat{\mathrm{B}}=\hat{0}$, that is $\hat{\mathrm{B}} v=0 \quad \forall v \in V_{1}$.
Proof: We consider three cases.

1. $\operatorname{dim} V_{1}<\operatorname{dim} V_{2}$

Due to

$$
\hat{\mathrm{B}} \mathrm{~T}_{1}(g) v=\mathrm{T}_{2}(g) \hat{\mathrm{B}} v,
$$

we have

$$
\hat{\mathrm{B}} v \in \operatorname{im} \hat{\mathrm{~B}} \quad \Rightarrow \quad \mathrm{~T}_{2}(g) \hat{\mathrm{B}} v \in \operatorname{im} \hat{\mathrm{~B}} \quad \forall \hat{\mathrm{~B}} v \in \operatorname{im} \hat{\mathrm{~B}}, g \in G .
$$

We see that im $\hat{\mathrm{B}}$ is a subspace invariant under $T_{2}$. Since $T_{2}$ is irreducible, im $\hat{\mathrm{B}}$ must be either $V_{2}$ or $\{0\}$. The first option is impossible, since it would lead us to $\operatorname{dim} \operatorname{im} \hat{B}=\operatorname{dim} V_{2}$. This contradicts

$$
\operatorname{dim} \operatorname{im} \hat{\mathrm{B}} \leq \operatorname{dim} V_{1}<\operatorname{dim} V_{2}
$$

Thus, we must accept the second option, im $\hat{B}=\{0\}$, and conclude that

$$
\hat{\mathrm{B}} v=0 \quad \forall v \in V_{1}
$$

2. $\operatorname{dim} V_{1}>\operatorname{dim} V_{2}$

Consider an element $k \in \operatorname{ker} \hat{\mathrm{~B}} \subseteq V_{1}$ so that

$$
\hat{\mathrm{B}} k=0 .
$$

By construction,

$$
\hat{\mathrm{B}} \mathrm{~T}_{1}(g) k=\mathrm{T}_{2}(g) \hat{\mathrm{B}} k=0 .
$$

Thus, we have $\mathrm{T}_{1}(g) k \in \operatorname{ker} \hat{\mathrm{~B}}$. Since

$$
k \in \operatorname{ker} B \quad \Rightarrow \quad \mathrm{~T}_{1}(g) k \in \operatorname{ker} \hat{\mathrm{~B}} \quad g \in G,
$$

we see that ker $\hat{B}$ is a subspace invariant under $T_{1}$. Since $T_{1}$ is irreducible, ker $\hat{B}$ is either $V_{1}$ or $\{0\}$. The second option is impossible, since due to $\operatorname{dim} V_{1}=\operatorname{dimim} \hat{\mathrm{B}}+$ $\operatorname{dim} \operatorname{ker} \hat{\mathrm{B}}$ and $\operatorname{dimim} \hat{\mathrm{B}} \leq \operatorname{dim} V_{2}$, $\operatorname{dim} \operatorname{ker} \hat{\mathrm{B}}$ cannot be zero. Thus, we must accept the first option, $\operatorname{ker} \hat{\mathrm{B}}=V_{1}$, and conclude

$$
\hat{\mathrm{B}} v=0 \quad \forall v \in V_{1}
$$

3. $\operatorname{dim} V_{1}=\operatorname{dim} V_{2}$

As in 2., ker $\hat{\mathrm{B}}$ is an invariant subspace, and must be either $V_{1}$ or $\{0\}$. In the first case, $\hat{\mathrm{B}}=0$. In the second case, $\hat{\mathrm{B}}$ is invertible and we have

$$
\begin{aligned}
\hat{\mathrm{B}} \mathrm{~T}_{1}(g) & =\mathrm{T}_{2}(g) \hat{\mathrm{B}}, \\
\mathrm{~T}_{1}(g) & =\mathrm{B}^{-1} \mathrm{~T}_{2}(g) \hat{\mathrm{B}} .
\end{aligned}
$$

That is, $T_{1}(g)$ and $T_{2}(g)$ are equivalent.

Theorem 2.7 (The Orthogonality Theorem) Let G be a finite group and let $\mathrm{D}_{\nu}$ and $\mathrm{D}_{\mu}$ be two irreducible representations. We have:

$$
\frac{\operatorname{dim} V_{\nu}}{\operatorname{ord} \mathrm{G}} \sum_{g \in \mathrm{G}} \mathrm{D}_{\nu}(g)_{q}^{i} \mathrm{D}_{\mu}\left(g^{-1}\right)^{p}{ }_{j}=\delta_{\nu \mu} \delta_{j}^{i} \delta^{p}{ }_{q}
$$

Proof: Let $\hat{\mathrm{A}}: V_{\nu} \rightarrow V_{\mu}$ be some transformation, and define

$$
\hat{\mathrm{B}}:=\sum_{g \in G} \mathrm{~T}_{\mu}(g) \hat{\mathrm{A}} \mathrm{~T}_{\nu}\left(g^{-1}\right) .
$$

Thus, for all $h \in \mathrm{G}$, and writing $g^{\prime}=h g$, we have

$$
\begin{aligned}
\mathrm{T}_{\mu}(h) \hat{\mathrm{B}} & =\sum_{g \in G} \mathrm{~T}_{\mu}(h) \mathrm{T}_{\mu}(g) \hat{\mathrm{A}} \mathrm{~T}_{\nu}\left(g^{-1}\right) \\
& =\sum_{g \in G} \mathrm{~T}_{\mu}(h g) \hat{\mathrm{A}_{\nu}\left(g^{-1}\right)} \\
& =\sum_{g^{\prime} \in G} \mathrm{~T}_{\mu}\left(g^{\prime}\right) \hat{\mathrm{A}_{\nu}}\left(g^{\prime-1} h\right) \\
& =\sum_{g^{\prime} \in G} \mathrm{~T}_{\mu}\left(g^{\prime}\right) \hat{\mathrm{A}_{\nu}\left(g^{\prime-1}\right) \mathrm{T}_{\nu}(h)} \\
\mathrm{T}_{\mu}(h) \hat{\mathrm{B}} & =\hat{\mathrm{B}} \mathrm{~T}_{\nu}(h)
\end{aligned}
$$

By Schur's First Lemma (Thm. 2.5), we have $\hat{B}=\lambda \mathbb{1}$. By Schur's Second Lemma (Thm. 2.6), $\hat{\lambda}=0$ unless $\mu=\nu$. Writing $\hat{\mathrm{B}}$ out in matrix form then gives us

$$
\mathbf{B}_{j}^{i}=\sum_{g \in G} \mathrm{D}_{\mu}(g)^{i}{ }_{m} \mathbf{A}^{m}{ }_{n} \mathrm{D}_{\nu}\left(g^{-1}\right)^{n}{ }_{j}=\lambda_{\mathbf{A}} \delta_{\nu \mu} \delta^{i}{ }_{j},
$$

where $\lambda_{\mathbf{A}}$ denotes the dependence of $\lambda$ on $\mathbf{A}$. This equation holds for all $\mathbf{A}$, so we choose $\mathbf{A}^{m}{ }_{n}=\delta^{m}{ }_{q} \delta^{p}{ }_{n}$ for $p, q$ fixed. Thus,

$$
\sum_{g \in \mathrm{G}} \mathrm{D}_{\mu}(g)^{i} \mathrm{D}_{\nu}\left(g^{-1}\right)^{p}{ }_{j}=\lambda_{p q} \delta_{\mu \nu} \delta^{i}{ }_{j}
$$

We solve for $\lambda_{p q}$ by setting $\nu=\mu$ and then taking the trace of both sides (i.e. setting $i=j$ ):

$$
\begin{aligned}
\sum_{g \in \mathrm{G}} \mathrm{D}_{\nu}(g)^{i}{ }_{q} \mathrm{D}_{\nu}\left(g^{-1}\right)^{p}{ }_{j} & =\lambda_{p q} \delta^{i}{ }_{j} \\
\sum_{g \in \mathrm{G}}\left(\mathrm{D}_{\nu}\left(g^{-1}\right) \mathrm{D}_{\nu}(g)\right)^{p}{ }_{q} & =\lambda_{p q} \operatorname{dim} V_{\nu} \\
\delta^{p}{ }_{q} \operatorname{ord} \mathrm{G} & =\lambda_{p q} \operatorname{dim} V_{\nu} \\
\lambda_{p q} & =\delta^{p}{ }_{q} \frac{\operatorname{ord} \mathrm{G}}{\operatorname{dim} V_{\nu}}
\end{aligned}
$$

Corollary 2.8 (Orthogonality of Characters) Let G be a finite group. We first define a "character product" by

$$
\left\langle\chi_{\mu}, \chi_{\nu}\right\rangle:=\frac{1}{\operatorname{ord} G} \sum_{g \in G} \chi_{\mu}(g) \chi_{\nu}\left(g^{-1}\right)
$$

Now let $\mathrm{D}_{\mu}, \mathrm{D}_{\nu}$ be two irreducible representations, and let $\chi_{\mu}, \chi_{\nu}$ denote their characters. We have

$$
\left\langle\chi_{\mu}, \chi_{\nu}\right\rangle=\delta_{\mu \nu}
$$

Consider a decomposition of a completely reducible representation

$$
\mathrm{D}=\mathrm{D}_{1} \oplus \mathrm{D}_{2} \oplus \cdots \oplus \mathrm{D}_{M}=\bigoplus_{\mu=1}^{M} \mathrm{D}_{\mu}
$$

It may be the case that some of the irreducible representations $\mathrm{D}_{\mu}$ are identical. For instance, if $\mathrm{D}_{3}=\mathrm{D}_{4}$, we could write

$$
\mathrm{D}_{1} \oplus \mathrm{D}_{2} \oplus 2 \mathrm{D}_{3} \oplus \cdots
$$

instead of

$$
\mathrm{D}_{1} \oplus \mathrm{D}_{2} \oplus \mathrm{D}_{3} \oplus \mathrm{D}_{4} \oplus \cdots
$$

Notice that when we write $2 \mathrm{D}_{\mu}$, we don't mean "multiply $\mathrm{D}_{\mu}$ by the number two." Instead, when we decompose a matrix $\mathrm{D}(g)$ into block diagonal form, we get two identical blocks $\mathrm{D}_{\mu}(g)$.
We can take advantage of this notation by slipping in a coefficient $a_{\nu} \in \mathbb{N}$ into the decomposition of D :

$$
\mathrm{D}=a_{1} \mathrm{D}_{1} \oplus a_{2} \mathrm{D}_{2} \oplus \cdots \oplus a_{N} \mathrm{D}_{N}=\bigoplus_{\nu=1}^{N} a_{\nu} \mathrm{D}_{\nu}
$$

The coefficient $a_{\nu}$ can be easily determined by the characters of the representations:

## Corollary 2.9 (Decomposition into Irreducible Components)

Let G be a finite group. Let D be a representation with character $\chi$. The decomposition of $D$ into irreducible representations is

$$
\mathrm{D}=\bigoplus_{\nu=1}^{N}\left\langle\chi, \chi_{\nu}\right\rangle \mathrm{D}_{\nu},
$$

where the index $\nu$ runs through all the inequivalent irreducible representations of G .
Corollary 2.10 (Clebsch-Gordan Decomposition) Let G be finite and let $\mathrm{D}_{\mu}, \mathrm{D}_{\nu}$ be two irreducible representations, with characters $\chi_{\mu}, \chi_{\nu}$. Their tensor product has the decomposition

$$
\mathrm{D}_{\mu} \otimes \mathrm{D}_{\nu}=\bigoplus_{\sigma=1}^{N}\left\langle\chi_{\sigma}, \chi_{\mu} \chi_{\nu}\right\rangle \mathrm{D}_{\sigma}
$$

where the index $\nu$ runs through all the inequivalent irreducible representations of G .

## Chapter 3

## Rotations

Recall that the group $C_{n}$ is generated by a single element $a$ which means "rotate by $\frac{2 \pi}{n}$." The mapping

$$
f(a)=e^{i \frac{2 \pi}{n}}
$$

associates abstract group elements to points of the complex plane. These points all lie on the unit circle.


Figure 3.1: Each point stands for a group element.
Since all these points represent rotations, we somehow expect the unit circle to contain all possible rotations.

### 3.1 The Lie Group SO(2)

We define the abstract group of proper rotations $S O(2)$ to contain all rotations about the origin of a two-dimensional plane, where a proper rotation denotes the absence of reflections. Unless otherwise specified, all rotations referred to are assumed to be proper. Unlike the cyclic groups $\mathrm{C}_{n}$, the group of rotations $S O(2)$ has infinitely many elements, which need to be specified using a continuous parameter $\varphi \in[0,2 \pi)$. We thus have group operation $\circ$, identity element $\mathbb{1}=\mathrm{id}$, and inverse element of a rotation by the same angle back. This group is furthermore abelian.

## Matrix Representations of $S O(2)$

One matrix representation R of $S O(2)$ under the standard basis $\left\{\mathbf{e}_{1}, \mathbf{e}_{2}\right\}$ is defined by

$$
\mathrm{R}: S O(2) \rightarrow \mathrm{GL}(2, \mathbb{R}) \quad \mathrm{R}(\varphi) \equiv \mathbf{R}_{\varphi}=\left[\begin{array}{cc}
\cos \varphi & -\sin \varphi \\
\sin \varphi & \cos \varphi
\end{array}\right]
$$

Since $\mathrm{R}(\varphi)=\mathbb{1}$ is fulfilled only for $\varphi=0$, which corresponds to the identity element, we see that ker $R=\{\mathbb{1}\}$. By Proposition 1.3, we see that $R$ is injective and therefore faithful. Notice also that $\mathrm{R}(\varphi)^{\mathrm{T}} \mathrm{R}(\varphi)=\mathbb{1}$ and $\operatorname{det} \mathrm{R}(\varphi)=1$; the using this faithful representation R , we can formulate an equivalent definition of $S O(2)$.
The group $\mathrm{SO}(2)$ is the group of all $2 \times 2$ orthogonal matrices $\mathbf{R}$

$$
\mathrm{SO}(2):=\left\{\mathbf{R} \in \mathrm{GL}(2, \mathbb{R}) \mid \mathbf{R}^{\top} \mathbf{R}=\mathbb{1} \quad \operatorname{det} \mathbf{R}=1\right\}
$$

Notice that since R is faithful, the abstract group $S O(2)$ of rotations and the group $\mathrm{SO}(2)$ of orthogonal matrices are identical.
The representation $\mathrm{R}: S O(2) \rightarrow \mathrm{GL}(2, \mathbb{R})$ is irreducible. However, the complex representation

$$
\mathrm{R}: S O(2) \rightarrow \mathrm{GL}(2, \mathbb{C}) \quad \mathrm{R}(\varphi)=\left[\begin{array}{cc}
e^{i \varphi} & 0 \\
0 & e^{-i \varphi}
\end{array}\right]
$$

is clearly reducible. It reduces into

$$
\begin{aligned}
& \mathrm{R}(\varphi)=\mathrm{U}(\varphi) \oplus \mathrm{U}(-\varphi) \\
& \mathrm{U}(\varphi)=e^{i \varphi}
\end{aligned}
$$

where $\mathrm{U}: S O(2) \rightarrow \mathrm{GL}(1, \mathbb{C})$ is irreducible and one-dimensional. The image of U is actually $\mathrm{U}(1)$, the group of all unitary $1 \times 1$ matrices (i.e. single numbers):

$$
\mathrm{U}(1):=\left\{\mathbf{M} \in \mathrm{GL}(1, \mathbb{C}) \mid \mathbf{M}^{\dagger} \mathbf{M}=\mathbb{1}\right\} .
$$

Since $U$ is injective on $U(1)$, it follows from the first isomorphism theorem (Thm. 1.5) that

$$
\mathrm{SO}(2) \cong \mathrm{U}(1)
$$

Notice that elements $U(\varphi)=e^{i \varphi}$ of $\mathrm{U}(1)$ lie along $S^{1}$, the unit circle in the complex plane

$$
S^{1}:=\{z \in \mathbb{C}| | z \mid=1\} \cong \mathrm{U}(1) .
$$

Thus, we have

$$
\mathrm{SO}(2) \cong S^{1}
$$

Just as the unit circle contains all points $e^{\frac{2 \pi \mathrm{i}}{n}}$, the group $S O(2)$ contains all groups $\mathrm{C}_{n}$. Notice that in the course of this script, we've shown a pretty long chain of isomorphisms:

$$
S O(2) \cong \mathrm{SO}(2) \cong \mathrm{U}(1) \cong S^{1} \cong[0,2 \pi) \cong \mathbb{R} / \mathbb{Z}
$$

## Manifolds and Lie groups

$\mathrm{SO}(2)$ is our first example of a continuous group, also known as a Lie group: a differentiable manifold whose elements satisfy the properties of a group. The precise definition of a manifold is too elaborate for our purposes, so we just briefly go over some of its key properties.
Manifolds have differing global and local structures. On a small scale, a manifold simply looks like a Euclidean space $\mathbb{R}^{n}$. However, on a large scale, a manifold can take on more interesting geometries. For instance, the manifold $S^{1}$ has the global structure of a circle, but the local structure of a straight line, $\mathbb{R}^{1}$.
A manifold $\mathcal{M}$ is called connected if it is not a union of two disjoint, non-empty, open sets. In a simply connected manifold, every closed path contained in $\mathcal{M}$ can be shrunk down to a single point. For instance, a torus is not simply connected.

## Infinitesimal Generators

Recall that $\mathrm{C}_{n}$ was generated by a single element $a$, which was a rotation about $\frac{2 \pi}{n}$. As we let $n$ go to infinity, we notice that this rotation gets smaller and smaller. In this way a very small rotation $\varphi \rightarrow 0$ can be said to generate the group $S O(2)$.
Let's look at the Taylor expansion of $\mathrm{R}(\varphi)$ around $\varphi=0$ :

$$
\mathrm{R}(\varphi)=\mathrm{R}(0)+\left.\varphi \cdot \frac{\mathrm{d}}{\mathrm{~d} \varphi}\right|_{\varphi=0} \mathrm{R}(\varphi)+\cdots
$$

Calculating the derivative

$$
\left.\frac{\mathrm{d}}{\mathrm{~d} \varphi}\right|_{\varphi=0} \mathrm{R}(\varphi)=\left.\frac{\mathrm{d}}{\mathrm{~d} \varphi}\right|_{\varphi=0}\left[\begin{array}{cc}
\cos \varphi & -\sin \varphi \\
\sin \varphi & \cos \varphi
\end{array}\right]=\left[\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right]=: \mathbf{X}
$$

gives

$$
R(\varphi)=\mathbb{1}+\varphi \mathbf{X}+\cdots
$$

The matrix $\varphi \mathbf{X}$ therefore induces an infinitesimal rotation $\varphi \rightarrow 0$. We call $\mathbf{X}$ an infinitesimal generator. We can achieve any finite rotation by "summing up" this infinitesimal generator $\mathbf{X}$ with the exponential function, which gives us the Taylor expansion of R.

$$
\exp (\varphi \mathbf{X})=\sum_{k=0}^{\infty} \frac{1}{k!} \varphi^{k} X^{k}=\mathbb{1}+\varphi \mathbf{X}+\cdots=R(\varphi)
$$

## Lie Algebras

In general, infinitesimal generators are actually not members of the Lie group they generate. Instead, they form a different kind of structure:
Let $\gamma$ be a curve in $G$ through $\mathbb{1}$. An infinitesimal generator $\mathbf{X}$ arises by differentiation of $\gamma(t)$ at the identity element:

$$
\mathbf{X}:=\left.\frac{\mathrm{d}}{\mathrm{~d} t}\right|_{t=0} \gamma(t)
$$

We thus consider $\mathbf{X}$ to be a tangent vector to $G$ at the identity element.
For higher-dimensional Lie groups, a family of curves $\gamma_{1}, \gamma_{2}, \ldots, \gamma_{n}$ will generate several tangent vectors $\mathbf{X}_{1}, \mathbf{X}_{2}, \ldots, \mathbf{X}_{n}$. The span of all tangent vectors is called the tangent space at $\mathbb{1}$.

Theorem 3.1 The tangent space at the identity element of a Lie group G forms a Lie algebra $\mathfrak{g}$.

As mentioned in Section 1.2, a Lie algebra is an algebra with a Lie-bracket as the algebra multiplication. For matrix groups and algebras, we use the commutator as algebra multiplication.

$$
[\mathbf{X}, \mathbf{Y}]=\mathbf{X Y}-\mathbf{Y X}
$$

Given a Lie algebra element X, we can "generate" a Lie group element by exponentiation:

$$
\exp (t \mathbf{X})=\gamma(t) \in \mathbf{G}
$$

### 3.2 The Lie Group SO(3)

Just as $S O(2)$ contained all rotations around the origin of a two-dimensional plane, we define $S O(3)$ to be the abstract Lie group of all rotations about the origin of a threedimensional Euclidean space $\mathbb{R}^{3}$. Similarly, the concrete Lie group $\mathrm{SO}(3)$ is the group of all $3 \times 3$ orthogonal matrices $\mathbf{R}$ :

$$
\mathrm{SO}(3):=\left\{\mathbf{R} \in \mathrm{GL}(3, \mathbb{R}) \mid \mathbf{R}^{\top} \mathbf{R}=\mathbb{1} \quad \operatorname{det} \mathbf{R}=1\right\}
$$

While $S O(2)$ required only one parameter $\varphi$ to characterize a rotation, $S O(3)$ requires three parameters.

## The Lie algebra so(3)

Let $\mathrm{R}: \mathbb{R}^{3} \rightarrow \mathrm{SO}(3)$ be some unknown parametrization. How can we calculate the Lie algebra so(3)? Since we don't have enough information to directly evaluate

$$
\mathbf{X}:=\left.\frac{\mathrm{d}}{\mathrm{~d} \boldsymbol{\varphi}}\right|_{\boldsymbol{\varphi}=0} R(\boldsymbol{\varphi}),
$$

we need to look at the Taylor expansion of $R(\boldsymbol{\varphi})$ for small rotations $\boldsymbol{\varphi} \rightarrow 0$

$$
R(\boldsymbol{\varphi})=\mathbb{1}+\boldsymbol{\varphi} \mathbf{X}+\cdots
$$

Since $R(\boldsymbol{\varphi}) \in \mathrm{SO}(3)$ is orthogonal, we have

$$
\begin{aligned}
& \mathbb{1}=\mathrm{R}(\boldsymbol{\varphi})^{\mathrm{T}} \mathrm{R}(\boldsymbol{\varphi})=(\mathbb{1}+\boldsymbol{\varphi} \mathbf{X}+\cdots)^{\mathrm{T}}(\mathbb{1}+\boldsymbol{\varphi} \mathbf{X}+\cdots) \\
& \mathbb{1}=\mathbb{1}+\boldsymbol{\varphi}\left(\mathrm{X}^{\mathrm{T}}+\mathbf{X}\right)+\cdots \\
& 0=\boldsymbol{\varphi}\left(\mathbf{X}^{\mathrm{T}}+\mathbf{X}\right)+\cdots
\end{aligned}
$$

where $\cdots$ stands for terms of order $\varphi^{2}$. Dividing by $\varphi$ and taking the limit $\varphi \rightarrow 0$ gives us

$$
\mathbf{X}^{\mathrm{T}}=-\mathbf{X}
$$

Any $\mathbf{X} \in$ so(3) must be antisymmetric, so we define

$$
\text { so(3) }:=\left\{\mathbf{X} \in \mathbb{R}^{3 \times 3} \mid \mathbf{X}^{\mathrm{T}}=-\mathbf{X}\right\}
$$

## Euler-Angle Parametrization

One of the many methods of parametrizing $S O(3)$ involves the use of Euler angles. Any rotation in $\mathbb{R}^{3}$ can be described as a series of rotations $\varphi_{1}, \varphi_{2}, \varphi_{3}$ about the $x$-, $y$-, and $z$-axes respectively. (Actually, the true Euler-Angle parametrization is quite different, but it eventually lead to this parametrization. See Tung, §7.1.) Thus, we can parametrize $S O(3)$ by three angles $\varphi_{1}, \varphi_{2}, \varphi_{3}$, and we have a faithful matrix representation

$$
\begin{aligned}
& \mathrm{R}: \mathbb{R}^{3} \rightarrow \mathrm{SO}(3) \quad \mathrm{R}(\boldsymbol{\varphi}) \equiv \mathrm{R}\left(\varphi_{1}, \varphi_{2}, \varphi_{3}\right)=\mathrm{R}_{3}\left(\varphi_{3}\right) \mathrm{R}_{2}\left(\varphi_{2}\right) \mathrm{R}_{1}\left(\varphi_{1}\right) \\
& \mathrm{R}_{1}\left(\varphi_{1}\right) \equiv \mathbf{R}_{x, \varphi_{1}}=\left[\begin{array}{ccc}
1 & 0 & 0 \\
0 & \cos \varphi_{1} & -\sin \varphi_{1} \\
0 & \sin \varphi_{1} & \cos \varphi_{1}
\end{array}\right] \\
& \mathrm{R}_{2}\left(\varphi_{2}\right) \equiv \mathbf{R}_{y, \varphi_{2}}=\left[\begin{array}{ccc}
\cos \varphi_{2} & 0 & \sin \varphi_{2} \\
0 & 1 & 0 \\
-\sin \varphi_{2} & 0 & \cos \varphi_{2}
\end{array}\right] \\
& \mathrm{R}_{3}\left(\varphi_{3}\right) \equiv \mathbf{R}_{z, \varphi_{3}}=\left[\begin{array}{ccc}
\cos \varphi_{3} & -\sin \varphi_{3} & 0 \\
\sin \varphi_{3} & \cos \varphi_{3} & 0 \\
0 & 0 & 1
\end{array}\right]
\end{aligned}
$$

Using this parametrization, we can easily calculate the infinitesimal generators:

$$
\begin{aligned}
& \mathbf{X}_{1}=\left.\frac{\mathrm{d}}{\mathrm{~d} \varphi_{1}}\right|_{\varphi_{1}=0} R_{1}\left(\varphi_{1}\right)=\left[\begin{array}{ccc}
0 & 0 & 0 \\
0 & 0 & -1 \\
0 & 1 & 0
\end{array}\right] \\
& \mathbf{X}_{2}=\left.\frac{\mathrm{d}}{\mathrm{~d} \varphi_{2}}\right|_{\varphi_{2}=0} R_{2}\left(\varphi_{2}\right)=\left[\begin{array}{ccc}
0 & 0 & 1 \\
0 & 0 & 0 \\
-1 & 0 & 0
\end{array}\right] \\
& \mathbf{X}_{3}=\left.\frac{\mathrm{d}}{\mathrm{~d} \varphi_{3}}\right|_{\varphi_{3}=0} R_{3}\left(\varphi_{3}\right)=\left[\begin{array}{ccc}
0 & -1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{array}\right] .
\end{aligned}
$$

These three infinitesimal generators form a basis of the Lie algebra so(3), which means that any $\mathbf{X} \in$ so(3) can be written as a linear combination of each $\mathbf{X}_{i}$

$$
\mathbf{X}=\xi^{i} \mathbf{X}_{i} \quad \xi^{i} \in \mathbb{R}
$$

The commutation relations of the $\mathbf{X}_{i}$ are

$$
\left[\mathbf{X}_{1}, \mathbf{X}_{2}\right]=\mathbf{X}_{3} \quad\left[\mathbf{X}_{2}, \mathbf{X}_{3}\right]=\mathbf{X}_{1} \quad\left[\mathbf{X}_{3}, \mathbf{X}_{1}\right]=\mathbf{X}_{2} .
$$

We can summarize all these relations using abstract index notation

$$
\begin{equation*}
\left(\mathbf{X}_{i}\right)^{j}{ }_{k}=-\epsilon_{i}{ }_{k}{ }_{k} \quad(\mathbf{X})_{j}^{i}=\epsilon_{j k}^{i} \xi^{k} \quad\left[\mathbf{X}_{i}, \mathbf{X}_{j}\right]=\epsilon_{i j}{ }^{k} \mathbf{X}_{k} \tag{3.1}
\end{equation*}
$$

Thus, the structure constants of so(3) are

$$
c_{j k}^{i}=\epsilon_{j k}^{i} .
$$

## Axis-Angle Parametrization

Another method of parametrizing $\mathrm{SO}(3)$ and so(3) specifies an angle $\varphi \in[0, \pi)$ about an axis of rotation determined by a unit vector $\mathbf{n}=n^{i} \mathbf{e}_{i}$.
The infinitesimal generator $\mathbf{X}_{\mathbf{n}} \in \operatorname{so}(3)$ is given by

$$
\begin{gathered}
\mathbf{X}_{\mathbf{n}} \equiv \mathbf{n} \cdot \mathbf{X}=n^{k} \mathbf{X}_{k} \\
\mathbf{X}_{\mathbf{n}}=\left[\begin{array}{ccc}
0 & -n^{3} & n^{2} \\
n^{3} & 0 & -n^{1} \\
-n^{2} & n^{1} & 0
\end{array}\right],
\end{gathered}
$$

and a finite rotation $\mathbf{R}_{\mathbf{n}, \varphi} \in \mathrm{SO}$ (3) by an angle $\varphi$ about $\mathbf{n}$ is given by

$$
\begin{aligned}
& \mathbf{R}_{\mathbf{n}, \varphi}=\exp \left(\varphi \mathbf{X}_{\mathbf{n}}\right)=\exp \left(\varphi n^{k} \mathbf{X}_{k}\right) \\
& \mathbf{R}_{\mathbf{n}, \varphi}=\left[\begin{array}{ccc}
\cos \varphi+\left(n^{1}\right)^{2}(1-\cos \varphi) & n^{1} n^{2}(1-\cos \varphi)-n^{3} \sin \varphi & n^{1} n^{3}(1-\cos \varphi)+n^{2} \sin \varphi \\
n^{1} n^{2}(1-\cos \varphi)+n^{3} \sin \varphi & \cos \varphi+\left(n^{2}\right)^{2}(1-\cos \varphi) & n^{2} n^{3}(1-\cos \varphi)-n^{1} \sin \varphi \\
n^{1} n^{3}(1-\cos \varphi)-n^{2} \sin \varphi & n^{2} n^{3}(1-\cos \varphi)+n^{1} \sin \varphi & \cos \varphi+\left(n^{3}\right)^{2}(1-\cos \varphi)
\end{array}\right]
\end{aligned}
$$

## Topology of SO(3)

While $\mathrm{SO}(2)$ is simply isometric to the unit circle $S^{1}, \mathrm{SO}(3)$ has a trickier topology. Consider the solid ball of radius $\pi$

$$
B_{\pi}:=\left\{\mathbf{x} \in \mathbb{R}^{3}| | \mathbf{r} \mid \leq \pi\right\} .
$$

The axis of rotation corresponds to the direction of $\mathbf{x} \in B_{\pi}$, and the angle of rotation corresponds to its length:

$$
\mathrm{n} \widehat{=} \frac{\mathrm{x}}{|\mathrm{x}|} \quad \varphi \widehat{=}|\mathrm{x}|
$$

The center $\mathbf{x}=0$ corresponds to the identity transformation. Rotation by a negative angle would correspond to a point along the same axis but across the origin. The only issue left is that a rotation $\varphi=\pi$ and $\varphi=-\pi$ are the same. This means that antipodal points on the surface of the ball are identical - a strange property indeed.
The space constructed in this way is called the real projective space $\mathbb{R} P^{3}$.

## Representation of SO(3) and so(3) on Wave Functions

We now turn our attention to a common representation used in quantum mechanics. Let $\mathcal{H} \equiv L^{2}\left(\mathbb{R}^{3}, \mathrm{~d} \mu\right)$ be the Hilbert space of quantum mechanical wave functions. We define the representation $\mathrm{T}: \mathrm{SO}(3) \rightarrow \mathrm{GL}(\mathcal{H})$ by

$$
\mathrm{T}(\mathbf{R}) \equiv \hat{\mathrm{R}} \quad(\hat{\mathrm{R}} \psi)(\mathbf{x}):=\psi\left(\mathbf{R}^{-1} \cdot \mathbf{x}\right) \quad \mathbf{R} \in \mathrm{SO}(3)
$$

This equation simply says that rotating a wave function in one direction is the same as rotating the coordinate axes in the other direction. Notice that this representation is faithful. Thus, we have an identical copy of $\mathrm{SO}(3)$ as a subgroup of $\mathrm{GL}(\mathcal{H})$.
We determine the Lie algebra representation $\mathrm{L}: \operatorname{so}(3) \rightarrow \mathrm{GL}(\mathcal{H})$ induced by T :

$$
\mathrm{L}\left(\mathbf{X}_{\mathbf{n}}\right) \equiv \hat{\mathrm{X}}_{\mathbf{n}} \quad\left(\hat{\mathrm{X}}_{n} \psi\right)(\mathbf{x}):=?
$$

First, we calculate $\hat{X}_{3}$ by using the definition of the infinitesimal generator

$$
\hat{\mathrm{X}}_{3}=\left.\frac{\mathrm{d}}{\mathrm{~d} \varphi}\right|_{\varphi=0} \hat{\mathrm{R}}_{3, \varphi}=\lim _{\varphi \rightarrow 0} \frac{\hat{\mathrm{R}}_{3, \varphi}-\hat{\mathrm{R}}_{3,0}}{\varphi}=\lim _{\varphi \rightarrow 0} \frac{\hat{\mathrm{R}}_{3, \varphi}-\mathbb{1}}{\varphi} .
$$

With $\mathbf{R}^{-1} \equiv \mathbf{R}_{3,-\varphi}$

$$
\begin{aligned}
\mathbf{R}_{3,-\varphi} \cdot \mathbf{x} & =\left[\begin{array}{ccc}
\cos \varphi & \sin \varphi & 0 \\
-\sin \varphi & \cos \varphi & 0 \\
0 & 0 & 1
\end{array}\right]\left[\begin{array}{l}
x \\
y \\
z
\end{array}\right] \\
& =\left[\begin{array}{c}
x \cos \varphi+y \sin \varphi \\
y \cos \varphi-x \sin \varphi \\
z
\end{array}\right]=\left[\begin{array}{c}
x+y \varphi+\cdots \\
y-x \varphi+\cdots \\
z
\end{array}\right]
\end{aligned}
$$

we calculate $\hat{\mathrm{R}}_{3, \varphi}$ :

$$
\begin{aligned}
\hat{\mathrm{R}}_{3, \varphi} \psi(\mathbf{x})= & \psi\left(\mathbf{R}_{3,-\varphi} \cdot \mathbf{x}\right) \\
= & \psi(x+y \varphi+\cdots, y-x \varphi+\cdots, z) \\
= & \left.\psi(x+y \varphi+\cdots, y-x \varphi+\cdots, z)\right|_{\varphi=0}+ \\
& +\left.\varphi \cdot \frac{\mathrm{d} \psi(x+y \varphi+\cdots, y-x \varphi+\cdots, z)}{\mathrm{d} \varphi}\right|_{\varphi=0}+\cdots \\
& =\psi(x, y, z)+\varphi \cdot\left(\frac{\partial \psi(x, y, z)}{\partial x} y-\frac{\partial \psi(x, y, z)}{\partial y} x\right)+\cdots \\
\Rightarrow \quad \hat{\mathrm{R}}_{3, \varphi}= & \hat{\mathbb{1}}+\varphi\left(y \frac{\partial}{\partial x}-x \frac{\partial}{\partial y}\right)+\cdots,
\end{aligned}
$$

where $\cdots$ refers to terms of order $\varphi^{2}$. Thus, we have for $\hat{\mathrm{X}}_{3}$ :

$$
\begin{aligned}
& \hat{\mathrm{X}}_{3}=\lim _{\varphi \rightarrow 0} \frac{\hat{\mathrm{R}}_{3, \varphi}-\mathbb{1}}{\varphi}=\lim _{\varphi \rightarrow 0} \frac{\hat{\mathbb{1}}+\varphi\left(y \frac{\partial}{\partial x}-x \frac{\partial}{\partial y}\right)+\cdots-\hat{\mathbb{1}}}{\varphi} \\
& \hat{\mathrm{X}}_{3}=y \frac{\partial}{\partial x}-x \frac{\partial}{\partial y}
\end{aligned}
$$

Similar calculations for rotations about the $x$ - and $y$ - axes give

$$
\hat{\mathrm{X}}_{1}=z \frac{\partial}{\partial y}-y \frac{\partial}{\partial z} \quad, \quad \hat{\mathrm{X}}_{2}=x \frac{\partial}{\partial z}-z \frac{\partial}{\partial x}
$$

We can thus summarize these results as

$$
\hat{\mathbf{X}}_{i}=-\epsilon_{i j}^{k} x^{j} \frac{\partial}{\partial x^{k}} \quad \hat{\mathbf{X}}=-\mathbf{x} \times \nabla
$$

where $\hat{\mathbf{X}}$ is the vector operator

$$
\hat{\mathbf{X}}:=\left(\hat{\mathrm{X}}_{1}, \hat{\mathrm{X}}_{2}, \hat{\mathrm{X}}_{3}\right)
$$

One can also calculate the commutation relations for the operators $\hat{\mathrm{X}}_{i}$, which are

$$
\left[\hat{\mathrm{X}}_{i}, \hat{\mathrm{X}}_{j}\right]=\epsilon_{i j}^{k} \hat{\mathrm{X}}_{k}
$$

Comparing $\hat{\mathbf{X}}$ to the orbital angular momentum operator $\hat{\mathbf{L}}=-\mathrm{i} \hbar \mathbf{x} \times \nabla$, we thus have

$$
\hat{\mathbf{L}}=\mathrm{i} \hbar \hat{\mathbf{X}}
$$

That is, the orbital angular momentum operator $\hat{\mathbf{L}}$ is identical to the operator $\hat{\mathbf{X}}$ that represents the Lie algebra so(3), up to a factor $\mathrm{i} \hbar$. We will investigate this relationship further on in the next chapter.

### 3.3 The Lie Group $\operatorname{SU}(2)$

Recall that we defined the special unitary group $\operatorname{SU}(2)$ as the group of all $2 \times 2$ unitary matrices with determinant 1 :

$$
\mathrm{SU}(2):=\left\{\mathbf{M} \in \mathrm{GL}(2, \mathbb{C}) \mid \mathbf{M}^{\dagger} \mathbf{M}=\mathbb{1} \quad \operatorname{det} \mathbf{M}=1\right\} .
$$

This group is a Lie group as well, and can be parametrized by two complex numbers $a, b$ as

$$
\mathrm{U}(a, b)=\left[\begin{array}{cc}
a & -\bar{b} \\
b & \bar{a}
\end{array}\right]
$$

The condition $\operatorname{det} \mathrm{U}(a, b)=1$ yields $|a|^{2}+|b|^{2}=1$. By writing

$$
a:=w+\mathrm{i} x \quad b:=y+\mathrm{i} z
$$

with $w, x, y, z \in \mathbb{R}$, we see that the condition $\operatorname{det} \mathrm{U}(a, b)=1$ delivers

$$
w^{2}+x^{2}+y^{2}+z^{2}=1
$$

Thus, the Lie Group $\operatorname{SU}(2)$ is isomorphic to $S^{3}$, the three-dimensional unit sphere in $\mathbb{R}^{4}$ :

$$
\mathrm{SU}(2) \cong S^{3}
$$

## The Lie algebra su(2)

To determine su(2), we recall that $\mathbf{A} \in \operatorname{su}(2)$ implies

$$
\exp (t \mathbf{A}) \in \operatorname{SU}(2)
$$

for all $t \in \mathbb{R}$. This leads us to
i) $\mathbb{1}=\exp (t \mathbf{A})^{\dagger} \exp (t \mathbf{A})$
ii) $1=\operatorname{det}(\exp (t \mathbf{A}))$
for all $t \in \mathbb{R}$. Differentiating condition i) at $t=0$ gives us

$$
\begin{aligned}
& 0=\left.\mathbf{A}^{\dagger} \exp (t \mathbf{A})^{\dagger} \exp (t \mathbf{A})\right|_{t=0}+\left.\exp (t \mathbf{A})^{\dagger} \mathbf{A} \exp (t \mathbf{A})\right|_{t=0} \\
& 0=\mathbf{A}^{\dagger}+\mathbf{A}
\end{aligned}
$$

Thus, $\mathbf{A}$ must be anti-hermitean: $\mathbf{A}^{\dagger}=-\mathbf{A}$. Condition ii) gives us

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

Differentiation yields that $\mathbf{A}$ must be traceless: $\operatorname{tr} \mathbf{A}=0$. We can therefore define

$$
\operatorname{su}(2):=\left\{\mathbf{A} \in \mathbb{C}^{2 \times 2} \mid \mathbf{A}^{\dagger}=-\mathbf{A}, \quad \operatorname{tr} \mathbf{A}=0\right\}
$$

A basis for $\mathrm{su}(2)$ is given by the matrices $\mathbf{s}_{i}$

$$
\begin{aligned}
\mathbf{s}_{1}=-\frac{1}{2}\left[\begin{array}{cc}
0 & \mathrm{i} \\
\mathrm{i} & 0
\end{array}\right] \quad \mathbf{s}_{2} & =-\frac{1}{2}\left[\begin{array}{cc}
0 & 1 \\
-1 & 0
\end{array}\right] \quad \mathbf{s}_{3}=-\frac{1}{2}\left[\begin{array}{cc}
\mathrm{i} & 0 \\
0 & -\mathrm{i}
\end{array}\right] \\
\operatorname{su}(2) & =\operatorname{span}\left\{\mathbf{s}_{1}, \mathbf{s}_{2}, \mathbf{s}_{3}\right\}
\end{aligned}
$$

so that any $\mathbf{A} \in \operatorname{su}(2)$ is given by

$$
a^{i} \mathbf{s}_{i} \quad a^{i} \in \mathbb{R}
$$

The commutation relations of $\mathbf{s}_{i}$ are given by

$$
\left[\mathbf{s}_{i}, \mathbf{s}_{j}\right]=\epsilon_{i j}{ }^{k} \mathbf{s}_{k}
$$

Thus, the structure constants of su(2) are

$$
c_{i j}^{k}=\epsilon_{i j}{ }^{k} .
$$

Notice that these are the same structure constants as the Lie algebra so(3). It should come as no surprise that these two algebras are actually identical. We will investigate this point later on in this chapter.

## The Vector space $H_{2}$

While the matrices $\mathbf{s}_{i}$ span (over $\mathbb{R}$ ) the vector space su(2) of $2 \times 2$ traceless anti-hermitean matrices, the Pauli matrices

$$
\boldsymbol{\sigma}_{1}=\left[\begin{array}{cc}
0 & 1 \\
1 & 0
\end{array}\right] \quad \boldsymbol{\sigma}_{2}=\left[\begin{array}{cc}
0 & -\mathrm{i} \\
\mathrm{i} & 0
\end{array}\right] \quad \boldsymbol{\sigma}_{3}=\left[\begin{array}{cc}
1 & 0 \\
0 & -1
\end{array}\right]
$$

span (also over $\mathbb{R}$ ) the space of $2 \times 2$ traceless hermitean matrices

$$
\begin{aligned}
& H_{2}:=\left\{\mathbf{M} \in \mathbb{C}^{2 \times 2} \mid \mathbf{M}^{\dagger}=\mathbf{M} \quad \operatorname{tr} \mathbf{M}=0\right\} \\
& H_{2}=\operatorname{span}\left\{\boldsymbol{\sigma}_{1}, \boldsymbol{\sigma}_{2}, \boldsymbol{\sigma}_{3}\right\} .
\end{aligned}
$$

We can regard $H_{2}$ as a real, three-dimensional vector space by considering the isomorphism $f: \mathbb{R}^{3} \rightarrow H_{2}$ defined by

$$
f\left(\mathbf{e}_{i}\right)=\sigma_{i}
$$

In this way, we can think of any $2 \times 2$ hermitean matrix $\boldsymbol{\sigma}$ as a real vector $\mathbf{x}$

$$
\mathbf{x}=x^{i} \mathbf{e}_{i} \quad \boldsymbol{\sigma}=f(\mathbf{x})=x^{i} \boldsymbol{\sigma}_{i}
$$

## Rotating vectors in $H_{2}$ and $\mathbb{R}^{3}$

Theorem 3.2 To rotate a vector $\sigma$ in $H_{2}$, we can use a $2 \times 2$ unitary matrix $\mathbf{U} \in \operatorname{SU}(2)$ in the following way:

$$
\boldsymbol{\sigma} \mapsto \mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger}
$$



Figure 3.2: Rotating a vector $\boldsymbol{\sigma}$

We prove this theorem by considering the scalar product

$$
\boldsymbol{\sigma} \cdot \boldsymbol{\varsigma}:=\frac{1}{2} \operatorname{tr}(\boldsymbol{\sigma} \boldsymbol{\varsigma})
$$

which defines the concept of an angle on $H_{2}$. We show

1. $\mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger}$ is still an element of $H_{2}$ :

$$
\left(\mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger}\right)^{\dagger}=\mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger} \quad \operatorname{tr}\left(\mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger}\right)=\operatorname{tr}\left(\boldsymbol{\sigma} \mathbf{U}^{\dagger} \mathbf{U}\right)=\operatorname{tr} \boldsymbol{\sigma}=0
$$

2. The mapping $\boldsymbol{\sigma} \mapsto \mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger}$ is isometric:

$$
\left(\mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger}\right) \cdot\left(\mathbf{U} \boldsymbol{\varsigma} \mathbf{U}^{\dagger}\right)=\frac{1}{2} \operatorname{tr}\left(\mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger} \mathbf{U} \boldsymbol{\varsigma} \mathbf{U}^{\dagger}\right)=\frac{1}{2} \operatorname{tr}(\boldsymbol{\sigma} \boldsymbol{\varsigma})=\boldsymbol{\sigma} \cdot \boldsymbol{\varsigma}
$$

## Axis-Angle Parametrization

We can parametrize $\operatorname{SU}(2)$ and $\operatorname{su}(2)$ by specifying an angle $\varphi \in[0, \pi]$ and a unit vector $\mathbf{n}=n^{i} \boldsymbol{\sigma}_{i}$ in $H_{2}$. An infinitesimal generator $\mathbf{s}_{n} \in \mathbf{s u}(2)$ is given by

$$
\begin{aligned}
& \mathbf{s}_{\mathbf{n}} \equiv \mathbf{n} \cdot \mathbf{s}=n^{i} \mathbf{s}_{i} \\
& \mathbf{s}_{\mathbf{n}}=-\frac{1}{2}\left[\begin{array}{cc}
\mathrm{i} n^{3} & \mathrm{i} n^{1}+n^{2} \\
\mathrm{i} n^{1}-n^{2} & -\mathrm{i} n^{3}
\end{array}\right],
\end{aligned}
$$

and a transformation $\mathbf{U}_{\mathbf{n}, \varphi} \in \operatorname{SU}(2)$ is given by

$$
\begin{aligned}
& \mathbf{U}_{\mathbf{n}, \varphi}=\exp \left(\varphi \mathbf{s}_{\mathbf{n}}\right)=\exp \left(\varphi n^{k} \mathbf{s}_{k}\right)=\exp \left(-\frac{\varphi \mathrm{i}}{2} \boldsymbol{\sigma}_{k}\right) \\
& \mathbf{U}_{\mathbf{n}, \varphi}=\mathbb{1} \cos \frac{\varphi}{2}-\mathrm{i} n^{k} \boldsymbol{\sigma}_{k} \sin \frac{\varphi}{2} \\
& \mathbf{U}_{\mathbf{n}, \varphi}=\left[\begin{array}{cc}
\cos \frac{\varphi}{2}-\mathrm{i} n^{3} \sin \frac{\varphi}{2} & -\sin \frac{\varphi}{2}\left(\mathrm{i} n^{1}+n^{2}\right) \\
-\sin \frac{\varphi}{2}\left(\mathrm{i} n^{1}-n^{2}\right) & \cos \frac{\varphi}{2}+\mathrm{i} n^{3} \sin \frac{\varphi}{2}
\end{array}\right],
\end{aligned}
$$

## Lie Group Homomorphism $\mathrm{SU}(2) \rightarrow \mathrm{SO}(3)$

Consider a rotation $\varphi$ around the " $z$-axis" of $H_{2}\left(\mathbf{n}=\boldsymbol{\sigma}_{3}\right)$ :

$$
\begin{aligned}
\mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger} & \equiv \mathbf{U}_{\mathbf{e}_{3, \varphi} \boldsymbol{\sigma}} \mathbf{U}_{\mathbf{e}_{3}, \varphi}^{\dagger}=\mathbf{U}_{\mathbf{e}_{3}, \varphi} \boldsymbol{\sigma} \mathbf{U}_{\mathbf{e}_{3},-\varphi} \\
& =\left[\begin{array}{cc}
e^{-\frac{i \varphi}{2}} & 0 \\
0 & e^{\frac{i \varphi}{2}}
\end{array}\right]\left(x \boldsymbol{\sigma}_{1}+y \boldsymbol{\sigma}_{2}+z \boldsymbol{\sigma}_{3}\right)\left[\begin{array}{cc}
e^{\frac{i \varphi}{2}} & 0 \\
0 & e^{-\frac{i \varphi}{2}}
\end{array}\right] \\
& =x\left[\begin{array}{cc}
0 & e^{-\mathrm{i} \varphi} \\
e^{\mathrm{i} \varphi} & 0
\end{array}\right]+\mathrm{i} y\left[\begin{array}{cc}
0 & -e^{-\mathrm{i} \varphi} \\
e^{\mathrm{i} \varphi} & 0
\end{array}\right]+z\left[\begin{array}{cc}
1 & 0 \\
0 & -1
\end{array}\right] \\
& =(x \cos \varphi-y \sin \varphi) \boldsymbol{\sigma}_{1}+(x \sin \varphi+y \cos \varphi) \boldsymbol{\sigma}_{2}+z \boldsymbol{\sigma}_{3}
\end{aligned}
$$

We see that rotating $\boldsymbol{\sigma} \in H_{2}$ around the $\boldsymbol{\sigma}_{3}$-axis using $\mathbf{U} \in \operatorname{SU}(2)$

$$
\boldsymbol{\sigma} \mapsto \mathbf{U} \boldsymbol{\sigma} \mathbf{U}^{\dagger} \quad\left[\begin{array}{l}
x \\
y \\
z
\end{array}\right] \mapsto\left[\begin{array}{c}
x \cos \varphi-y \sin \varphi \\
x \sin \varphi+y \cos \varphi \\
z
\end{array}\right]
$$

is identical to rotating $\mathbf{x} \in \mathbb{R}^{3}$ around the $\mathbf{e}_{3}$-axis using $\mathbf{R} \in \mathrm{SO}$ (3)

$$
\mathbf{x} \mapsto \mathbf{R} \mathbf{x} \quad\left[\begin{array}{l}
x \\
y \\
z
\end{array}\right] \mapsto\left[\begin{array}{c}
x \cos \varphi-y \sin \varphi \\
x \sin \varphi+y \cos \varphi \\
z
\end{array}\right]
$$

We see that the groups $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$, which may not seem similar at first sight, are actually quite intimately related. We express this relationship concretely in the form of a homomorphism $\Phi: \operatorname{SU}(2) \rightarrow \mathrm{SO}(3)$

$$
\mathbf{R} \mathbf{x}=\Phi(\mathbf{U}) \mathbf{x}:=f^{-1}\left(\mathbf{U} f(\mathbf{x}) \mathbf{U}^{\dagger}\right)
$$

where $f$ is the isomorphism $f(\mathbf{x})=\boldsymbol{\sigma}$ between $H_{2}$ and $\mathbb{R}^{3}$. (In this case $f$ is also called an intertwiner.)


Figure 3.3: A long way and a short way of producing the same rotation.

## Theorem 3.3 We have

1. $\Phi$ is surjective
2. $\operatorname{ker} \Phi=\{\mathbb{1},-\mathbb{1}\}=: Z_{2}$

Because of $i i$ ), we see that $\Phi$ is not injective. In fact

$$
\Phi(-\mathbf{U})=\Phi(-\mathbb{1}) \Phi(\mathbf{U})=\Phi(\mathbf{U})=\mathbf{R}
$$

That is, every element $\mathbf{R} \in \mathrm{SO}(3)$ is mapped by two elements of $\operatorname{SU}(2)$ :

$$
\Phi^{-1}(\mathbf{R})=\{\mathbf{U},-\mathbf{U}\}
$$

Thus, we say that $\Phi$ is a two-fold covering.
Another result of $i i$ ), as well as theorem 1.5 is

$$
\mathrm{SU}(2) / \mathrm{Z}_{2} \cong \mathrm{SO}(3)
$$

Although $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$ are not globally isomorphic, they are locally isomorphic. This concept manifests itself in their Lie algebras.

## Lie Algebra Isomorphism su(2) $\rightarrow$ so(3)

Recall that the Lie algebras su(2) and so(3) have the same structure constants. This suggests that these algebras are isomorphic. We can actually show this by using the group homomorphism $\Phi: \mathrm{SU}(2) \rightarrow \mathrm{SO}(3)$ to construct an isomorphism $\phi: \operatorname{su}(2) \rightarrow \mathrm{so}(3)$. First, we express the homomorphism $\Phi$ using abstract index notation

$$
\begin{aligned}
f(\mathbf{R x}) & =f(\Phi(\mathbf{U}) \mathbf{x})=\mathbf{U} f(\mathbf{x}) \mathbf{U}^{\dagger} \\
R_{j}^{i} x^{j} \boldsymbol{\sigma}_{i} & =\mathbf{U} x^{j} \sigma_{j} \mathbf{U}^{\dagger} .
\end{aligned}
$$

We differentiate this equation at $t=0$, noting that $\left.\frac{\mathrm{d}}{\mathrm{d} t}\right|_{t=0} \mathbf{R}=\mathbf{X} \in \operatorname{so}(3),\left.\mathbf{U}\right|_{t=0}=\mathbb{1}$, and $\left.\frac{\mathrm{d}}{\mathrm{d} t}\right|_{t=0} \mathbf{U}=\mathbf{A}=a^{i} \mathbf{s}_{i} \in \operatorname{su}(2)$, we have

$$
\begin{aligned}
\mathbf{X}_{j}^{i} x^{j} \boldsymbol{\sigma}_{i} & =\mathbf{A} x^{j} \boldsymbol{\sigma}_{j}+x^{j} \boldsymbol{\sigma}_{j} \mathbf{A}^{\dagger} \\
x^{j} \mathbf{X}_{j}^{i} \boldsymbol{\sigma}_{i} & =x^{j}\left(\mathbf{A} \boldsymbol{\sigma}_{j}-\boldsymbol{\sigma}_{j} \mathbf{A}\right) \\
& =x^{j}\left[\mathbf{A}, \boldsymbol{\sigma}_{j}\right]=x^{j}\left[a^{i} \mathbf{s}_{i}, \boldsymbol{\sigma}_{j}\right]=-\frac{i}{2} x^{j} a^{i}\left[\boldsymbol{\sigma}_{i}, \boldsymbol{\sigma}_{j}\right] \\
& =-\frac{i}{2} x^{j} a^{i}\left(2 \mathrm{i} \epsilon_{i j}{ }^{k} \boldsymbol{\sigma}_{k}\right) \\
& =x^{j} a^{i} \epsilon_{i j}{ }^{k} \boldsymbol{\sigma}_{k}=x^{j}\left(-a^{k} \epsilon_{k j}{ }^{i}\right) \boldsymbol{\sigma}_{i} \\
& =: x^{j} \phi(\mathbf{A})_{j}{ }^{i} \boldsymbol{\sigma}_{i} .
\end{aligned}
$$

We have thus found the isomorphism $\phi: \operatorname{su}(2) \rightarrow \mathbf{s o}(3)$

$$
\mathbf{X}_{j}^{i}=\phi(\mathbf{A})^{i}{ }_{j}=a^{k} \epsilon^{i}{ }_{j k}
$$

Thus, the Lie algebras su(2) and so(3) are isomorphic. Because Lie algebras are locally identical to their corresponding Lie groups, we can say that $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$ are locally isomorphic.
Notice also that these Lie algebras have the same structure constants as the Lie algebra $\left(\mathbb{R}^{3}, \times\right)$ introduced in chapter 1 . These algebras are thus isomorphic.

## Representations of $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$

If we are given some representation $\mathrm{D}: \mathrm{SO}(3) \rightarrow \mathrm{GL}(V)$, we can uniquely define a representation $\tilde{\mathrm{D}}: \mathrm{SU}(2) \rightarrow \mathrm{GL}(V)$ by

$$
\tilde{\mathrm{D}}(\mathbf{U}):=\mathrm{D}(\Phi(\mathbf{U}))
$$

However, if we this time let $\tilde{\mathrm{D}}$ be given, then the definition

$$
\mathrm{D}(\mathbf{R}):=\tilde{\mathrm{D}}\left(\Phi^{-1}(\mathbf{R})\right)
$$

is not unique, since $\Phi^{-1}(\mathbf{R})= \pm \mathbf{U}$, which gives us

$$
\mathrm{D}(\mathbf{R})=\tilde{\mathrm{D}}( \pm \mathbf{U})=\tilde{\mathrm{D}}( \pm \mathbb{1}) \tilde{\mathrm{D}}(\mathbf{U})
$$

and we end up with two different kinds of representations $\tilde{D}$ :

1. $\tilde{D}(-\mathbb{1})=+\mathbb{1}$

This kind of representation leads us to a unique, singly-defined representation D of SO(3)

$$
\mathrm{D}(\mathbf{R})=\tilde{\mathrm{D}}( \pm \mathbb{1}) \tilde{\mathrm{D}}(\mathbf{U})=+\mathbb{1} \tilde{\mathrm{D}}(\mathbf{U})
$$

This kind of representation is used for orbital angular momentum $\hat{\mathbf{L}}$.
2. $\tilde{D}(-\mathbb{1})=-\mathbb{1}$

This kind of representation leads us to a doubly-defined "representation" D of SO(3)

$$
\mathrm{D}(\mathbf{R})=\tilde{\mathrm{D}}( \pm \mathbb{1}) \tilde{\mathrm{D}}(\mathbf{U})= \pm \mathbb{1} \tilde{\mathrm{D}}(\mathbf{U})= \pm \tilde{\mathrm{D}}(\mathbf{U})
$$

This kind of representation is used for spin angular momentum $\hat{\mathbf{S}}$ and is referred to as a spinor representation of $\mathrm{SO}(3)$. Notice, however, that it is not a representation in the mathematical sense.

## Chapter 4

## Angular Momentum

Recall from quantum mechanics that any angular momentum operator, whether orbital $\hat{\mathbf{J}}=\hat{\mathbf{L}}$, spin $\hat{\mathbf{J}}=\hat{\mathbf{S}}$, or combined $\hat{\mathbf{J}}=\sum \hat{\mathbf{L}}+\sum \hat{\mathbf{S}}$ must satisfy the following commutation relations:

$$
\left[\hat{\mathrm{J}}_{i}, \hat{J}_{j}\right]=\mathrm{i} \hbar \epsilon_{i j k} \hat{\mathrm{~J}}_{k} .
$$

If we replace $\hat{\mathrm{J}}_{i}$ by some hypothetical operator $\mathrm{i} \hbar \hat{\mathrm{X}}_{i}$, we see that

$$
\begin{aligned}
{\left[(\mathrm{i} \hbar) \hat{\mathrm{X}}_{i},(\mathrm{i} \hbar) \hat{\mathrm{X}}_{j}\right] } & =\mathrm{i} \hbar \epsilon_{i j k}(\mathrm{i} \hbar) \hat{\mathrm{X}}_{k} \\
{\left[\hat{\mathrm{X}}_{i}, \hat{\mathrm{X}}_{j}\right] } & =\epsilon_{i j k} \hat{\mathrm{X}}_{k}
\end{aligned}
$$

That is, these operators $\hat{\mathrm{X}}_{i}$ have the structure constants $\epsilon_{i j k}$ and thus form an su(2) $\cong \mathrm{so}(3)$ Lie algebra. Thus, all angular momentum operators $\hat{\mathrm{J}}_{i}$ are simply operators $\hat{\mathrm{X}}_{i}$ which are given by a representation of $\operatorname{su}(2)$ or so(3), up to a factor $\mathrm{i} \hbar$.
We have already encountered a situation like this in the last chapter. Recall that the orbital angular momentum operator $\hat{\mathbf{L}}$ was given by

$$
\hat{\mathrm{L}}_{i}=\mathrm{i} \hbar \mathrm{~L}\left(\mathbf{X}_{i}\right) \equiv \mathrm{i} \hbar \hat{\mathrm{X}}_{i},
$$

where $L$ is an operator representation of so(3). The operators $\hat{\mathrm{X}}_{i}$, which generate an so(3) Lie-algebra, are equal to the orbital angular momentum operators $\hat{\mathrm{L}}_{i}$ up to a factor $\mathrm{i} \hbar$. This script follows the convention

$$
\mathrm{QM}=\mathrm{i} \hbar \mathrm{GT} .
$$

This factor $\mathrm{i} \hbar$ provides the link between group-theoretical operators, which are produced by representations of Lie-groups and Lie-algebras, and quantum-mechanical operators, which correspond to physically meaningful quantities.
(Beware of varying conventions! Some physics-oriented texts split up this factor i $\hbar$ and conceal the complex number i in some distant definition. For instance, Jones defines the infinitesimal generator as: $\mathbf{X}:=-\left.\mathrm{i} \frac{\mathrm{d}}{\mathrm{d} \varphi}\right|_{\varphi=0} \mathrm{R}(\varphi)$, which convinces the reader that $\mathrm{QM}=\hbar \mathrm{GT}$. Cornwell uses an entirely different convention, $\mathrm{QM}=\mathrm{i} \hbar \mathrm{GT}$, which results in left-handed rotations. The convention used in this script matches that of Penrose.)

### 4.1 Spin Angular Momentum

The goal of this section is to find the irreducible representations of su(2). Just as so(3) produced orbital angular momentum operators, we will soon see that su(2) produces spin angular momentum operators.

## The Lie-Algebra sl( $2, \mathbb{C}$ )

One small but vital detail that is often left out in many physics-oriented texts is the consideration of the Lie-algebra sl$(2, \mathbb{C})$ :

$$
s \mid(2, \mathbb{C}):=\left\{\mathbf{M} \in \mathbb{C}^{2 \times 2} \mid \operatorname{tr} \mathbf{M}=0\right\},
$$

which is the Lie-algebra of the Lie-group $\operatorname{SL}(2, \mathbb{C})$. It turns out that $\mathbf{s l}(2, \mathbb{C})$ can be generated by complex linear combinations of the matrices $\mathbf{s}_{i}$ :

$$
\begin{aligned}
\mathrm{sl}(2, \mathbb{C}) & =\left\{\mathbf{M} \in \mathbb{C}^{2 \times 2} \mid \mathbf{M}=\alpha^{i} \mathbf{s}_{i} \quad \alpha^{i} \in \mathbb{C}\right\} \\
& \equiv \operatorname{span}_{\mathbb{C}}\left\{\mathbf{s}_{i}\right\}_{i=1}^{3}
\end{aligned}
$$

Comparing this to su(2)

$$
\operatorname{su}(2)=\operatorname{span}_{\mathbb{R}}\left\{\mathbf{s}_{i}\right\}_{i=1}^{3},
$$

we clearly see that $\operatorname{su}(2) \subset \operatorname{sl}(2, \mathbb{C})$. Thus, any representation T of $\mathrm{sl}(2, \mathbb{C})$ is also a representation of $\operatorname{su}(2)$; all we need to do is restrict the domain of T .

## The Casimir Operator $\hat{\mathbf{J}}^{2}$

From the angular momentum operators $\hat{\mathrm{J}}_{i}$ we can construct the operator

$$
\hat{\mathbf{J}}^{2}:=\hat{\mathrm{J}}_{i} \hat{\mathrm{~J}}_{i} \equiv \hat{\mathrm{~J}}_{1}^{2}+\hat{\mathrm{J}}_{2}^{2}+\hat{\mathrm{J}}_{3}^{2},
$$

which commutes with all operators $\hat{J}_{i}$ :

$$
\begin{aligned}
{\left[\hat{\mathrm{J}}_{i}, \hat{\mathbf{J}}^{2}\right] } & =\sum_{j=1}^{3}\left[\hat{\mathrm{~J}}_{i}, \hat{\mathrm{~J}}_{j} \hat{\mathrm{~J}}_{j}\right]=\sum_{j=1}^{3}\left[\hat{\mathrm{~J}}_{i}, \hat{\mathrm{~J}}_{j}\right] \hat{\mathrm{J}}_{j}+\hat{\mathrm{J}}_{j}\left[\hat{\mathrm{~J}}_{i}, \hat{\mathrm{~J}}_{j}\right] \\
& =\sum_{j=1}^{3} \mathrm{i} \hbar \epsilon_{i j}{ }^{k} \hat{\mathrm{~J}}_{k} \hat{\mathrm{~J}}_{j}+\mathrm{i} \hbar \epsilon_{i j}{ }^{k} \hat{\mathrm{~J}}_{j} \hat{\mathrm{~J}}_{k} \\
& =0
\end{aligned}
$$

This property makes $\hat{\mathbf{J}}^{2}$ a Casimir operator. Although it is not an element of the Lie algebra $\mathrm{sl}(2, \mathbb{C})$, we can still use it to help us construct an irreducible basis.
Since we're interested in finding an irreducible representation of $\boldsymbol{s l}(2, \mathbb{C})$, we apply Schur's first lemma (Theorem 2.5; the theorem carries over literally to algebras) and notice that

$$
\hat{\mathbf{J}}^{2}=\lambda \hat{\mathbb{1}} .
$$

That is, $\hat{\mathbf{J}}^{2}$ has only one eigenvalue $\lambda \in \mathbb{R}$, which needs to be determined.
Since $\hat{\mathbf{J}}^{2}$ and $\hat{\mathrm{J}}_{3}$ commute, we have a system $\left\{\hat{\mathbf{J}}^{2}, \hat{\mathrm{~J}}_{3}\right\}$ of commuting observables. There is a basis of common eigenvectors, which we denote by $|\lambda m\rangle$, where $m$ stands for the different eigenvalues of $\widehat{J}_{3}$ :

$$
\hat{\mathbf{J}}^{2}|\lambda m\rangle=\hbar^{2} \lambda|\lambda m\rangle \quad \hat{\mathrm{J}}_{3}|\lambda m\rangle=\hbar m|\lambda m\rangle \ldots
$$

## The Eigenvalue Ladder

In order to determine the eigenvalues $\lambda$ and $m$, we define the ladder operators

$$
\hat{\mathrm{J}}_{+}:=\hat{\mathrm{J}}_{1}+\mathrm{i} \hat{\mathrm{~J}}_{2} \quad \hat{\mathrm{~J}}_{-}:=\hat{\mathrm{J}}_{1}-\mathrm{i} \hat{\mathrm{~J}}_{2}
$$

which satisfy the following commutation relations:

$$
\left[\hat{\mathrm{J}}_{+}, \hat{\mathrm{J}}_{-}\right]=2 \hbar \hat{\mathrm{~J}}_{3} \quad\left[\hat{\mathrm{~J}}_{3}, \hat{\mathrm{~J}}_{ \pm}\right]= \pm \hbar \hat{\mathrm{J}}_{ \pm}
$$

(Notice that the ladder operators are constructed by complex linear combinations of the basis operators $\hat{J}_{i}$-we cannot avoid $\mathrm{sl}(2, \mathbb{C})$.) The reason for the term "ladder operator" can be seen if we consider some eigenvalue $m$ of $\hat{J}_{3}$ :

$$
\hat{\mathrm{J}}_{3}|\lambda m\rangle=\hbar m|\lambda m\rangle
$$

Operating both sides by $\hat{J}_{+}$gives

$$
\begin{aligned}
\hat{\mathrm{J}}_{+} \hat{\mathrm{J}}_{3}|\lambda m\rangle & =\hbar m \hat{\mathrm{~J}}_{+}|\lambda m\rangle \\
\left(\hat{\mathrm{J}}_{3} \hat{\mathrm{~J}}_{+}-\hbar \hat{\mathrm{J}}_{+}\right)|\lambda m\rangle & =\hbar m \hat{\mathrm{~J}}_{+}|\lambda m\rangle \\
\hat{\mathrm{J}}_{3} \hat{\mathrm{~J}}_{+}|\lambda m\rangle & =\hbar(m+1) \hat{\mathrm{J}}_{+}|\lambda m\rangle \ldots
\end{aligned}
$$

While $|\lambda m\rangle$ is an eigenvector with eigenvalue $m$, we see that $\hat{J}_{+}|\lambda m\rangle$ is also an eigenvector, but with eigenvalue $m+1$. Similarly, $\hat{\mathrm{J}}_{-}|\lambda m\rangle$ is also an eigenvector with eigenvalue $m-1$. We can thus imagine a ladder that represents the eigenvectors and eigenvalues of $\hat{J}_{3}$, where the ladder operators $\hat{\mathrm{J}}_{ \pm}$let us "climb" up and down.
Since $V$ is finite-dimensional, we can expect this ladder to be of finite size. Thus, there exists an eigenvector

$$
|\lambda M\rangle
$$

with the largest eigenvalue possible $M$, which is represented by the topmost rung of the ladder. This eigenvector is sometimes referred to as the highest-weight vector. If we try to operate with $\hat{J}_{+}$again, we get

$$
\hat{\mathrm{J}}_{+}|\lambda M\rangle=0
$$

since otherwise, we would get an eigenvalue $M+1$, which would be higher than the maximum eigenvalue, leading to contradiction. Operating the above equation with $\hat{\mathrm{J}}_{-}$ and using

$$
\hat{\mathrm{J}}_{-} \hat{\mathrm{J}}_{+}=\left(\hat{\mathrm{J}}_{1}-\mathrm{i} \hat{\mathrm{~J}}_{2}\right)\left(\hat{\mathrm{J}}_{1}+\mathrm{i} \hat{\mathrm{~J}}_{2}\right)=\hat{\mathrm{J}}_{1}^{2}+\hat{\mathrm{J}}_{2}^{2}+\mathrm{i}\left[\hat{\mathrm{~J}}_{1}, \hat{\mathrm{~J}}_{2}\right]=\hat{\mathbf{J}}-\hat{\mathrm{J}}_{3}^{2}-\hbar \hat{\mathrm{J}}_{3}
$$



Figure 4.1: Climbing up and down with the ladder operators


Figure 4.2: The highest-weight vector, sitting at the top of the ladder
gives:

$$
\begin{aligned}
\hat{\mathrm{J}}_{-} \hat{\mathrm{J}}_{+}|\lambda M\rangle & =0 \\
\left(\hat{\mathbf{J}}^{2}-\hat{\mathrm{J}}_{3}^{2}-\hbar \hat{\mathrm{J}}_{3}\right)|\lambda M\rangle & =0 \\
\left(\lambda-M^{2}-M\right)|\lambda M\rangle & =0 \\
\Rightarrow \quad M(M+1) & =\lambda .
\end{aligned}
$$

We see that the sole eigenvalue $\lambda$ of the Casimir operator $\hat{\mathbf{J}}^{2}$ is given by $M(M+1)$.
From the highest-weight vector $|\lambda M\rangle$, we climb down to the lowest-weight vector $|\lambda \mu\rangle$ by $N$ applications of $\hat{\mathrm{J}}_{-}$.

$$
\hat{\mathrm{J}}_{-}^{N}|\lambda M\rangle=c_{N}|\lambda \mu\rangle \quad \mu=M-N \quad N \in \mathbb{N}_{0}
$$

where $c_{N}$ is some constant depending on conventions. Just like with the highest-weight vector, applying $\hat{J}_{-}$on the lowest-weight vector gives us zero. We apply $\hat{J}_{+}$and see that

$$
\begin{aligned}
\hat{\mathrm{J}}_{-}|\lambda \mu\rangle & =0 \\
\hat{\mathrm{~J}}_{+} \hat{\mathrm{J}}_{-}|\lambda \mu\rangle & =0 \\
\left(\hat{\mathbf{J}}^{2}-\hat{\mathrm{J}}_{3}^{2}+\hbar \hat{\mathrm{J}}_{3}\right)|\lambda \mu\rangle & =0 \\
\left(\lambda-\mu^{2}+\mu\right)|\lambda \mu\rangle & =0 \\
\Rightarrow \quad M(M+1)-(M-N)^{2}+(M-N) & =0 \\
M+2 N M-N^{2}+M-N & =0 \\
2 M(N+1) & =N(N+1)
\end{aligned}
$$



Figure 4.3: Climbing down to the lowest-weight vector
Since $N \in \mathbb{N}_{0}$, we see that the maximum eigenvalue $M=\frac{N}{2}$ can only take on non-negative integer and half-integer values:

$$
M=\frac{N}{2}=0, \frac{1}{2}, 1, \frac{3}{2}, \ldots
$$

The minimum eigenvalue $\mu$ is given by

$$
\mu=M-N=-\frac{N}{2}=0,-\frac{1}{2},-1,-\frac{3}{2}, \ldots
$$

We thus see that the eigenvalues are entirely determined by $M$, which we will instead denote by $j$, which is known as the spin number. All possible eigenvalues $m$ are thus given by

$$
\begin{gathered}
-j \leq m \leq j \quad j=0,-\frac{1}{2},-1,-\frac{3}{2}, \ldots \\
\frac{/ / / / / / / / / /}{}|j j\rangle \\
\frac{\vdots}{/ / / / / / / / / /}|j-j\rangle
\end{gathered}
$$

Figure 4.4: The complete eigenvalue ladder

## The irreducible representations of $\mathrm{su}(2)$

Notice that the representations given in the following theorems are of $\mathbf{s l}(2, \mathbb{C})$, and consequently of su(2) by restricting the domain, as mentioned above.

Theorem 4.1 The irreducible representations of $\mathrm{sl}(2, \mathbb{C})$ are given by a number $j=$ $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. They have dimension $2 j+1$ and there exists a basis $|j m\rangle, m=-j, \ldots, j$ such that

$$
\begin{aligned}
\hat{\mathrm{J}}_{3}|j m\rangle & =\hbar m|j m\rangle \\
\hat{\mathrm{J}}_{ \pm}|j m\rangle & =\hbar \sqrt{j(j+1)-m(m \pm 1)}|j, m \pm 1\rangle \\
\hat{\mathbf{J}}^{2}|j m\rangle & =\hbar^{2} j(j+1)|j m\rangle
\end{aligned}
$$

where $\hat{\mathrm{J}}_{1}=\frac{1}{2}\left(\hat{\mathrm{~J}}_{+}+\hat{\mathrm{J}}_{-}\right)$and $\hat{\mathrm{J}}_{2}=\frac{1}{2 \mathrm{i}}\left(\hat{\mathrm{J}}_{+}-\hat{\mathrm{J}}_{-}\right)$. The number $j$ is called the spin number of the representation, and the carrier space $\mathcal{S}$ is called the spinor space.

Proof: The representations are indeed irreducible; it suffices to notice that

$$
\mathcal{S}=\operatorname{span}\left\{\left(\hat{\mathrm{J}}_{-}\right)^{k}|j j\rangle\right\}_{k=0}^{2 j+1}
$$

Conversely, let an irreducible representation be given. Since $\hat{\mathbf{J}}^{2}$ and $\hat{J}_{3}$ commute, there exists a basis of common eigenvectors. The argument given in the last subsection gives the basis $\{|j m\rangle\}$.
We now calculate the coefficients of $\hat{J}_{ \pm}$:

$$
\hat{\mathrm{J}}_{ \pm}|j m\rangle=c_{ \pm}|j, m \pm 1\rangle
$$

We choose a scalar product so that $|j m\rangle$ are orthonormal. Then

$$
\begin{aligned}
\left|c_{ \pm}\right|^{2} & =\langle j m| \hat{\mathrm{J}}_{\mp} \hat{\mathrm{J}}_{ \pm}|m j\rangle \\
& =\langle j m|\left(\hat{\mathbf{J}}^{2}-\hat{\mathrm{J}}_{3}^{2} \mp \hbar J_{3}\right)|m j\rangle \\
& =\hbar^{2}(j(j+1)-m(m \pm 1))
\end{aligned}
$$

The value $c_{ \pm}$can be determined uniquely, except for some phase factor $e^{i \alpha}$ which we set to one. (This is in accordance with the Condon-Shortley convention.) Thus

$$
c_{ \pm}=\hbar \sqrt{j(j+1)-m(m \pm 1)}
$$

The above representations could be denoted explicitly by $\mathrm{T}^{j}$ so that

$$
\hat{\mathrm{J}}_{i}=\mathrm{i} \hbar \mathrm{~T}^{j}\left(\mathbf{s}_{i}\right) \quad i=1,2,3 \quad j=0, \frac{1}{2}, 1, \frac{3}{2}, \ldots
$$

However, the exact form of the right-hand side of this equation is subject to convention. More important are the operators $\hat{\mathrm{J}}_{i}$, which are independent of convention and appear in every text as given in Theorem (4.1).

Theorem 4.2 The irreducible matrix representations $\mathrm{D}^{j}$ of $\mathrm{su}(2)$ are given by the representations with spin $j$ under the basis $\{|j m\rangle\}=\left\{\mathbf{e}_{m-j+1}\right\}$. The matrices

$$
\mathbf{J}_{i}=\mathrm{i} \hbar \mathrm{D}^{j}\left(\mathbf{s}_{i}\right)
$$

are explicitly given by

$$
\begin{aligned}
& \left(\mathbf{J}_{1}\right)^{m^{\prime}}{ }_{m}=\frac{\hbar}{2} \sqrt{j(j+1)-m(m+1)} \delta^{m^{\prime}}{ }_{m+1}+\frac{\hbar}{2} \sqrt{j(j+1)-m(m-1)} \delta^{m^{\prime}}{ }_{m-1} \\
& \left(\mathbf{J}_{2}\right)^{m^{\prime}}{ }_{m}=\frac{\hbar}{2 \mathrm{i}} \sqrt{j(j+1)-m(m+1)} \delta^{m^{\prime}}{ }_{m+1}-\frac{\hbar}{2 \mathrm{i}} \sqrt{j(j+1)-m(m-1)} \delta^{m^{\prime}}{ }_{m-1} \\
& \left(\mathbf{J}_{3}\right)^{m^{\prime}}{ }_{m}=\hbar m \delta^{m^{\prime}}{ }_{m}
\end{aligned}
$$

We can write down these matrices explicitly for two simple cases.
Case 1: $j=\frac{1}{2}$ This case corresponds to a particle of spin $\frac{1}{2}$. The spin angular momentum operator $\hat{\mathbf{J}} \equiv \hat{\mathbf{S}}$ is given by the irreducible representation $\mathrm{T}^{\frac{1}{2}}$ of $\operatorname{su}(2)$ :

$$
\hat{\mathrm{S}}_{i}=\mathrm{i} \hbar \mathrm{~T}^{\frac{1}{2}}\left(\mathbf{s}_{i}\right) \ldots
$$

The spinor space $\mathcal{S}$ is of dimension 2 and spanned by the "spin up" and "spin down" eigenstates of the spin operator $\hat{\mathrm{S}}_{3}$.

$$
\begin{aligned}
& \left|\frac{1}{2} \frac{1}{2}\right\rangle \equiv|+\rangle=\mathbf{e}_{1}=\left[\begin{array}{l}
1 \\
0
\end{array}\right] \\
& \left|\frac{1}{2}-\frac{1}{2}\right\rangle \equiv|-\rangle=\mathbf{e}_{2}=\left[\begin{array}{l}
0 \\
1
\end{array}\right] \\
& \frac{/ / / / / / / 1 / /}{}|+\rangle \\
& \overline{/ / / / / / / / / /}|-\rangle
\end{aligned}
$$

Figure 4.5: The spinor space of dimension two
Under this basis, the spin matrices $\mathbf{S}_{i}=\mathrm{i} \hbar \mathrm{D}^{\frac{1}{2}}\left(\mathbf{s}_{i}\right)$ take on the form

$$
\begin{aligned}
& \mathbf{S}_{1}=\frac{\hbar}{2}\left[\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right]=\frac{\hbar}{2} \boldsymbol{\sigma}_{1} \\
& \mathbf{S}_{2}=\frac{\hbar}{2}\left[\begin{array}{cc}
0 & -i \\
i & 0
\end{array}\right]=\frac{\hbar}{2} \boldsymbol{\sigma}_{2} \\
& \mathbf{S}_{3}=\frac{\hbar}{2}\left[\begin{array}{cc}
1 & 0 \\
0 & -1
\end{array}\right]=\frac{\hbar}{2} \boldsymbol{\sigma}_{3},
\end{aligned}
$$

Notice that they are identical to the Pauli matrices up to the factor $\hbar / 2$. Furthermore, we see that the two-dimensional matrix representation $D^{\frac{1}{2}}$ is simply the identity mapping.

$$
\mathrm{D}^{\frac{1}{2}}\left(\mathbf{s}_{i}\right)=\frac{1}{\mathrm{i} \hbar} \mathbf{S}_{i}=-\frac{\mathrm{i}}{2} \sigma_{i}=\mathbf{s}_{i} .
$$

With $\mathbf{s}_{\mathbf{n}}=n^{k} \mathbf{S}_{k}$ as the infinitesimal generator of the axis-angle parametrization, we have

$$
\mathrm{D}^{\frac{1}{2}}\left(\mathbf{s}_{\mathbf{n}}\right)=\mathrm{s}_{\mathbf{n}} .
$$

Case 2: $j=1$ This case corresponds to a particle of spin 1 . The spin angular momentum operator is given by

$$
\hat{\mathrm{S}}_{i}=\mathrm{i} \hbar \mathrm{~T}^{1}\left(\mathbf{s}_{i}\right) .
$$

The spinor space $\mathcal{S}$ is of dimension three and spanned by the triplet of eigenstates of the operator $\hat{S}_{3}$.

$$
\begin{gathered}
|11\rangle=\mathbf{e}_{1}=\left[\begin{array}{l}
1 \\
0 \\
0
\end{array}\right] \\
|10\rangle=\mathbf{e}_{2}=\left[\begin{array}{l}
0 \\
1 \\
0
\end{array}\right] \\
|1-1\rangle=\mathbf{e}_{3}=\left[\begin{array}{l}
0 \\
0 \\
1
\end{array}\right] \\
/ / / / / / / / / / /
\end{gathered}|11\rangle \begin{array}{ll}
1 & 0\rangle \\
\hline 1 & |1-1\rangle \\
\hline / / / / / / / / / /
\end{array}
$$

Figure 4.6: The spinor space of dimension three
Under this basis, we get the following spin matrices $S_{i}$ :

$$
\begin{aligned}
& \mathbf{S}_{1}=\frac{\hbar \sqrt{2}}{2}\left[\begin{array}{lll}
0 & 1 & 0 \\
1 & 0 & 1 \\
0 & 1 & 0
\end{array}\right] \\
& \mathbf{S}_{2}=\frac{\hbar \sqrt{2}}{2 \mathrm{i}}\left[\begin{array}{ccc}
0 & 1 & 0 \\
-1 & 0 & 1 \\
0 & -1 & 0
\end{array}\right] \\
& \mathbf{S}_{3}=\hbar\left[\begin{array}{ccc}
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & -1
\end{array}\right] .
\end{aligned}
$$

A few simple calculations show us that $\mathbf{S}_{1}, \mathbf{S}_{2}, \mathbf{S}_{3}$ are equivalent to the infinitesimal generators $\mathbf{X}_{1}, \mathbf{X}_{2}, \mathbf{X}_{3}$ of $\mathrm{SO}(3)$.

### 4.2 Orbital Angular Momentum and Rotational Symmetry

We can use the irreducible representations $D^{j}$ of su(2) to obtain irreducible representations $\mathrm{D}^{j}$ of $\mathrm{SO}(3)$. We do this in two steps

## Step 1: From su(2) to $\operatorname{SU}(2)$

We can obtain a representation $\tilde{\mathrm{D}}^{j}$ of $\mathrm{SU}(2)$ by exponentiating the representation $\mathrm{D}^{j}$ of su(2)

$$
\tilde{\mathrm{D}}^{j}(\mathbf{U})=\tilde{\mathrm{D}}^{j}(\exp \varphi \mathbf{s})=\exp \left(\varphi \mathrm{D}^{j}(\mathbf{s})\right)
$$

This works because $\operatorname{SU}(2)$ is simply connected. We can solve the above exponential for the two special cases from the last subsection:
Case 1: $\quad j=\frac{1}{2}$ Just as $D^{\frac{1}{2}}(\mathbf{s})=\mathbf{s}$, we see that $\tilde{D}^{\frac{1}{2}}(\mathbf{U})=\mathbf{U}$ :

$$
\begin{aligned}
\tilde{D}^{\frac{1}{2}}\left(\mathbf{U}_{\mathbf{n}, \varphi}\right) & =\exp \left(\varphi n^{k} \mathrm{D}^{\frac{1}{2}}\left(\mathbf{s}_{k}\right)\right)=\exp \left(\varphi n^{k} \mathbf{s}_{k}\right) \\
& =\mathbf{U}_{\mathbf{n}, \varphi}=\left[\begin{array}{cc}
\cos \frac{\varphi}{2}-\mathrm{i} n^{3} \sin \frac{\varphi}{2} & -\sin \frac{\varphi}{2}\left(\mathrm{i} n^{1}+n^{2}\right) \\
-\sin \frac{\varphi}{2}\left(\mathrm{i} n^{1}-n^{2}\right) & \cos \frac{\varphi}{2}+\mathrm{i} n^{3} \sin \frac{\varphi}{2}
\end{array}\right] .
\end{aligned}
$$

Case 2: $\quad j=1$ The three-dimensional representation $\tilde{\mathrm{D}}^{j}$ of $\mathrm{SU}(2)$ is given by

$$
\tilde{\mathrm{D}}^{1}\left(\mathbf{U}_{\mathbf{n}, \varphi}\right)=\exp \left(\varphi n^{k} \mathrm{D}^{1}\left(\mathbf{s}_{k}\right)\right) .
$$

Exponentiating these matrices proves to be a daunting task, and the results themselves are unimportant. However, we can give an explicit result for a rotation $\varphi$ around the z-axis:

$$
\begin{aligned}
\tilde{\mathrm{D}}^{1}\left(\mathbf{U}_{\mathbf{e}_{3}, \varphi}\right) & =\exp \left(\varphi \mathrm{D}^{1}\left(\mathbf{s}_{3}\right)\right) \\
& =\exp \left(-\mathrm{i} \varphi\left[\begin{array}{ccc}
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & -1
\end{array}\right]\right)=\left[\begin{array}{ccc}
e^{-\mathrm{i} \varphi} & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & e^{\mathrm{i} \varphi}
\end{array}\right] .
\end{aligned}
$$

This is the diagonalization of the matrix

$$
\mathbf{R}_{\mathbf{e}_{3, \varphi}}=\left[\begin{array}{ccc}
\cos \varphi & -\sin \varphi & 0 \\
\sin \varphi & \cos \varphi & 0 \\
0 & 0 & 1
\end{array}\right]
$$

Note that we consider the three-dimensional complex representation, hence $\mathbf{R}_{\mathbf{e}_{3, \varphi}}$ can be diagonalized. Thus, the three-dimensional matrix representation $\tilde{D}^{1}$ of $\mathrm{SU}(2)$ is equivalent to the three-dimensional matrix representation R of $\mathrm{SO}(3)$ obtained by extending the carrier space $\mathbb{R}^{3}$ of the identical representation of $\mathrm{SO}(3)$ to $\mathbb{C}^{3}$. This follows also by comparing the traces:

$$
\chi\left(\mathbf{U}_{\mathbf{e}_{3}, \varphi}\right) \equiv \operatorname{tr} \quad \tilde{\mathrm{D}}^{1}\left(\mathbf{U}_{\mathbf{e}_{3}, \varphi}\right)=1+2 \cos \varphi, \quad \chi\left(\mathbf{R}_{\mathbf{e}_{3}, \varphi}\right) \equiv \operatorname{tr} \quad \mathbf{R}_{\mathbf{n}, \varphi}=1+2 \cos \varphi
$$

## Step 2: From $\operatorname{SU}(2)$ to $\mathrm{SO}(3)$

As described in the last chapter, we can define a representation D of $\mathrm{SO}(3)$ by means of a representation $\tilde{D}$ of $S U(2)$ by

$$
\mathrm{D}(\mathbf{R}):=\tilde{\mathrm{D}}\left(\Phi^{-1}(\mathbf{R})\right)
$$

However, there are two different types of representations $\tilde{\mathrm{D}}$ of $\mathrm{SU}(2)$

1. $\tilde{D}(-\mathbb{1})=+\mathbb{1} \quad$ This kind of representation defines a single-valued representation D of SO(3) by

$$
\mathrm{D}(\mathbf{R}):=\tilde{\mathrm{D}}(\mathbf{U})
$$

2. $\tilde{D}(-\mathbb{1})=-\mathbb{1} \quad$ This kind of representation does not give a well-defined representation D of $\mathrm{SO}(3)$, since

$$
\mathrm{D}(\mathbf{R}):= \pm \tilde{\mathrm{D}}(\mathbf{U})
$$

Case 1: $\quad j=\frac{1}{2}$ We see that

$$
\tilde{D}^{\frac{1}{2}}(-\mathbb{1})=\tilde{\mathrm{D}}^{\frac{1}{2}}\left(\mathbf{U}_{\mathbf{e}_{3}}, 2 \pi\right)=-\mathbb{1}
$$

By the above arguments, we see that we can't rely on $\tilde{D}^{\frac{1}{2}}$ to produce a well-defined representation $D^{\frac{1}{2}}$. Hence, it is referred to as a "double-valued" representation.
Case 2: $\quad j=1$ We see that

$$
\tilde{\mathrm{D}}^{1}(-\mathbb{1})=\tilde{\mathrm{D}}^{1}\left(\mathbf{U}_{\mathbf{e}_{3}, 2 \pi}\right)=\mathbb{1}
$$

Therefore, we can define a representation $\mathrm{D}^{1}$ of $\mathrm{SO}(3)$ by

$$
\mathrm{D}^{1}(\mathbf{R}):=\tilde{\mathrm{D}}^{1}(\mathbf{U})
$$

We just saw that this representation is equivalent to the standard matrix representation R of $\mathrm{SO}(3)$, with carrier space extended from $\mathbb{R}^{3}$ to $\mathbb{C}^{3}$ :

$$
\mathrm{D}^{1}\left(\mathbf{R}_{\mathbf{n}, \varphi}\right)=\exp \left(\varphi \mathrm{D}^{1}\left(\mathbf{s}_{\mathbf{n}}\right)\right) \quad \sim \quad \mathrm{R}(\mathbf{n}, \varphi) \equiv \mathbf{R}_{\mathbf{n}, \varphi}
$$

Theorem 4.3 A matrix representation $\tilde{\mathrm{D}}^{j}$ of $\mathrm{SU}(2)$ defines a single-valued matrix representation $\mathrm{D}^{j}$ of $\mathrm{SO}(3)$ if the spin number $j$ is an integer, that is $j=0,1,2, \ldots$.
Proof: Under the axis-angle parametrization of $\operatorname{SU}(2)$, we see that

$$
-\mathbb{1}=\mathbf{U}_{\mathbf{e}_{3}, 2 \pi} .
$$

Therefore, with $\mathrm{D}^{j}\left(\mathbf{s}_{3}\right)^{m^{\prime}}{ }_{m}=\frac{1}{\mathrm{i} \hbar}\left(\mathbf{J}_{3}\right)^{m^{\prime}}{ }_{m}=-\mathrm{i} m \delta^{m^{\prime}}{ }_{m}$, we have

$$
\begin{aligned}
\tilde{\mathrm{D}}^{j}(-\mathbb{1}) & =\tilde{\mathrm{D}}^{j}\left(\mathbf{U}_{\mathbf{e}_{3}, 2 \pi}\right)=\tilde{\mathrm{D}}^{j}\left(\exp \left(2 \pi \mathrm{~s}_{3}\right)\right)=\exp \left(2 \pi \mathrm{D}^{j}\left(\mathbf{s}_{3}\right)\right) \\
& =\exp -2 \pi \mathrm{i}\left[\begin{array}{llll}
j & & & 0 \\
& j-1 & & \\
& & \ddots & \\
0 & & & -j
\end{array}\right]=\left[\begin{array}{llll}
e^{2 \pi \mathrm{i} j} & & \\
& e^{2 \pi \mathrm{i}(j-1)} & & \\
& & \ddots & \\
0 & & & e^{-2 \pi \mathrm{i} j}
\end{array}\right] \\
& =(-1)^{2 j} \mathbb{1}
\end{aligned}
$$

As just described, a representation $D$ is well-defined by $\mathrm{D}(\mathbf{R}):=\tilde{\mathrm{D}}(\mathbf{U})$ if $\tilde{\mathrm{D}}(-\mathbb{1})=\mathbb{1}$. This is the case for $\tilde{\mathrm{D}}^{j}$ if $j=0,1,2, \ldots$

## Basis of Irreducible Functions of SO(3)

Recall the representation T of $\mathrm{SO}(3)$ on the space $L^{2}\left(\mathbb{R}^{3}, \mathrm{~d} \mu\right)$ given by

$$
\mathrm{T}(\mathbf{R}) \equiv \hat{\mathrm{R}} \quad \hat{\mathrm{R}} \psi(\mathbf{x}):=\psi\left(\mathbf{R}^{-1} \cdot \mathbf{x}\right)
$$

This representation can be decomposed into a direct sum of irreducible representations:

$$
\mathrm{T}=\bigoplus_{l=0}^{\infty} \mathrm{T}^{l}
$$

Reducing T into irreducible components requires finding subspaces $S_{l}$ of the carrier space $L^{2}\left(\mathbb{R}^{3}, \mathrm{~d} \mu\right)$ that are invariant under $\mathrm{T}^{l}$. The bases of these subspaces are given by the eigenvectors $|l m\rangle$ expressed as spatial wavefunctions:

$$
\psi_{l m}(\mathbf{x})=\langle\mathbf{x} \mid l m\rangle
$$

In order to calculate these basis functions $\psi_{l m}$, we take a detour and consider the representations of so(3).

Theorem 4.4 Let T be a representation of a Lie Group G , and L be a representation of the Lie algebra $\mathfrak{g}$, both acting on a carrier space $V$. A subspace $S \subseteq V$ is invariant under T if and only if it is invariant under L .

Therefore, we can instead reduce the representation L of so(3)

$$
\mathrm{L}\left(\mathbf{X}_{i}\right)=-(\mathrm{x} \times \nabla)_{i}
$$

into a direct sum of irreducible representations $\mathrm{T}^{l}$ of $\operatorname{su}(2) \cong \operatorname{so}(3)$ given in Theorem 4.1:

$$
\mathrm{L}=\bigoplus_{l=0}^{\infty} \mathrm{T}^{l}
$$

We begin by expressing the operators $\mathrm{L}\left(\mathbf{X}_{i}\right)$ in spherical coordinates

$$
\begin{aligned}
& \hat{\mathrm{L}}_{1}=\mathrm{i} \hbar \mathrm{~L}\left(\mathbf{X}_{1}\right)=\mathrm{i} \hbar\left(\sin \varphi \frac{\partial}{\partial \theta}+\cot \theta \cos \varphi \frac{\partial}{\partial \varphi}\right) \\
& \hat{\mathrm{L}}_{2}=\mathrm{i} \hbar \mathrm{~L}\left(\mathbf{X}_{2}\right)=\mathrm{i} \hbar\left(-\cos \varphi \frac{\partial}{\partial \theta}+\cot \theta \sin \varphi \frac{\partial}{\partial \varphi}\right) \\
& \hat{\mathrm{L}}_{3}=\mathrm{i} \hbar \mathrm{~L}\left(\mathbf{X}_{3}\right)=-\mathrm{i} \hbar \frac{\partial}{\partial \varphi}
\end{aligned}
$$

By Theorem 4.1, we know how the operators $\hat{\mathrm{L}}_{i}=\mathrm{i} \hbar \mathrm{L}\left(\mathbf{X}_{i}\right)=\mathrm{i} \hbar \bigoplus T^{l}\left(\mathbf{s}_{i}\right)$ act on the eigenvectors $|l m\rangle$

$$
\left.\begin{array}{rl}
\mathrm{i} \hbar \mathrm{~T}^{l}\left(\mathbf{s}_{3}\right)|l m\rangle & =\hat{\mathrm{L}}_{3}|l m\rangle
\end{array}=\hbar m|l m\rangle, \begin{array}{rl}
\mathrm{i} \hbar\left(\mathrm{~T}^{l}\left(\mathbf{s}_{1}\right)-\mathrm{iT}^{l}\left(\mathbf{s}_{2}\right)\right)|l m\rangle & =\hat{\mathrm{L}}_{-}|l m\rangle
\end{array}=\hbar \sqrt{l(l+1)-m(m-1)}|l m-1\rangle\right)
$$

Formally, this is achieved by

$$
\langle\mathbf{x}| \hat{\mathrm{L}}_{i}|l m\rangle=\langle\mathbf{x}| \hat{\mathrm{L}}_{i} \int \mathrm{~d}^{3} x^{\prime}\left|\mathbf{x}^{\prime}\right\rangle\left\langle\mathbf{x}^{\prime} \mid l m\right\rangle=\langle\mathbf{x}| \hat{\mathrm{L}}_{i}|\mathbf{x}\rangle\langle\mathbf{x} \mid l m\rangle \equiv \hat{\mathrm{L}}_{i} \psi_{l m}(\mathbf{x}),
$$

where the new $\hat{\mathrm{L}}_{i}$ are the corresponding operators acting on wavefunctions $\psi_{l m}$ rather than abstract eigenvectors $|l m\rangle$. Expressing these equations as spatial wavefunctions gives

$$
\begin{aligned}
\hat{\mathrm{L}}_{3} \psi_{l m} & =\hbar m \psi_{l m} \\
\hat{\mathrm{~L}}_{-} \psi_{l m} & =\hbar \sqrt{l(l+1)-m(m-1)} \psi_{l(m-1)} \\
\hat{\mathbf{L}}^{2} \psi_{l m} & =\hbar l(l+1) \psi_{l m}
\end{aligned}
$$

We now take the operators $\hat{\mathrm{L}}_{i}$ in their spherical-coordinate form and plug them into the above equations:

$$
\begin{align*}
-\mathrm{i} \frac{\partial}{\partial \varphi} \psi_{l m} & =m \psi_{l m}  \tag{4.1}\\
e^{-\mathrm{i} \varphi}\left(-\frac{\partial}{\partial \theta}+\mathrm{i} \cot \theta \frac{\partial}{\partial \varphi}\right) \psi_{l m} & =\sqrt{l(l+1)-m(m-1)} \psi_{l(m-1)}  \tag{4.2}\\
-\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial}{\partial \theta}\right)+\frac{1}{\sin ^{2} \theta} \frac{\partial^{2}}{\partial \varphi^{2}}\right] \psi_{l m} & =l(l+1) \psi_{l m} . \tag{4.3}
\end{align*}
$$

Equation 1 has the solution

$$
\psi_{l m}(r, \theta, \varphi)=e^{\mathrm{i} m \varphi} F(r, \theta) .
$$

Putting this into Equation 2 gives

$$
-\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial}{\partial \theta}\right)+\frac{m^{2}}{\sin ^{2} \vartheta}\right] F(r, \theta)=l(l+1) F(r, \theta),
$$

which is the associated Legendre equation, whose solution is

$$
F(r, \theta)=P_{l m}(\cos \theta) f(r),
$$

where $f$ is some function of $r$ only, and $P_{l m}(\xi)$ is the associated Legendre function. We now define the functions

$$
Y_{l m}(\theta, \varphi)=(-1)^{m} \sqrt{\frac{(2 j+1)(j-m)!}{4 \pi(j+m)!}} e^{i m \varphi} P_{l m}(\cos \theta)
$$

which are called the spherical harmonics. (The actual steps leading to the spherical harmonics $Y_{l m}$ and the general solution $\psi_{l m}$ are long and cumbersome. A satisfying derivation can be found in any decent book on quantum mechanics and/or differential equations.) The solution to equations $1-3$ is then

$$
\psi_{l m}(r, \theta, \varphi)=f(r) Y_{l m}(\theta, \varphi)
$$

## Partial Wave Decomposition

The above equation reflects the fact that we can split up a wavefunction $\psi \in L^{2}\left(\mathbb{R}^{3}, \mathrm{~d} \mu\right)$ into radial and angular parts:

$$
\begin{aligned}
L^{2}\left(\mathbb{R}^{3}, \mathrm{~d} \mu\right) & =L^{2}\left(\mathbb{R}_{+}, \mathrm{d} r\right) \otimes L^{2}\left(S^{2}, \sin \theta \mathrm{~d} \theta \mathrm{~d} \varphi\right) \\
\psi_{l m} & =f \otimes Y_{l m}
\end{aligned}
$$

We also see that $Y_{l m}$ serves as an orthonormal basis for $L^{2}\left(S^{2}, \sin \theta \mathrm{~d} \theta \mathrm{~d} \varphi\right)$, that is, any $\phi \in L^{2}\left(S^{2}\right)$ can be written as a linear combination of $Y_{l m}$

$$
\phi=\alpha^{l m} Y_{l m}
$$

and thus, any $\psi \in L^{2}\left(\mathbb{R}^{3}\right)$ can be written as

$$
\begin{aligned}
\psi & =f \otimes \phi=f \otimes \alpha^{l m} Y_{l m} \\
\psi(r, \theta, \varphi) & =\alpha^{l m} f(r) Y_{l m}(\theta, \varphi)
\end{aligned}
$$

To be precise, the representation $\mathrm{T}: \mathrm{SO}(3) \rightarrow \mathrm{GL}\left(L^{2}\left(\mathbb{R}^{3}\right)\right)$ can also be split up as above:

$$
\mathrm{T}(\mathbf{R}) \equiv \hat{\mathrm{R}}^{\prime}=\mathbb{1} \otimes \hat{\mathrm{R}}
$$

with $\hat{\mathrm{R}}^{\prime}: L^{2}\left(\mathbb{R}^{3}\right) \rightarrow L^{2}\left(\mathbb{R}^{3}\right)$ and $\hat{\mathrm{R}}: L^{2}\left(S^{2}\right) \rightarrow L^{2}\left(S^{2}\right)$. This reflects the fact that a rotation operator $\hat{\mathrm{R}}$ leaves the radial part of a wavefunction alone.
A subspace $S_{l} \subset L^{2}\left(S^{2}\right)$ invariant under $T^{l}$ is thus given by

$$
S_{l}=\operatorname{span}\left\{Y_{l m}\right\}_{m=-l}^{l}
$$

For instance, the subspace

$$
S_{1}=\operatorname{span}\left\{\sin \theta e^{-\mathrm{i} \varphi}, \cos \theta, \sin \theta e^{\mathrm{i} \varphi}\right\}
$$

is invariant under L and $T_{l}$ so that for all angular momentum operators $\hat{\mathrm{L}}_{1}, \hat{\mathrm{~L}}_{2}, \hat{\mathrm{~L}}_{3}$, we have

$$
\phi \in S_{1} \quad \Rightarrow \quad \hat{\mathrm{~L}}_{i} \phi \in S_{1} .
$$

## Spherically Symmetric Potentials

Consider the Hamilton operator

$$
\hat{\mathrm{H}}=-\hbar^{2} \Delta+V(r),
$$

where the potential $V$ is spherically symmetric and only depends on the radius $r$. The following commutation relations then hold:

$$
\left[\hat{\mathrm{H}}, \hat{\mathbf{L}}^{2}\right]=0 \quad\left[\hat{\mathrm{H}}, \hat{\mathrm{~L}}_{3}\right]=0
$$

making $\hat{\mathrm{H}}, \hat{\mathbf{L}}^{2}, \hat{\mathrm{~L}}_{3}$ a system of commuting observables, and thus the state space $\mathcal{H}=$ $L^{2}\left(\mathbb{R}^{3}, \mathrm{~d} \mu\right)$ is spanned by a basis of eigenvectors $|\alpha l m\rangle$ where

$$
\langle\mathbf{x} \mid \alpha l m\rangle=f_{\alpha}(r) Y_{l m}(\theta, \varphi)
$$

Here, $\alpha$ stands for the radial quantum numbers (e.g. energy).

## Chapter 5

## Adding Angular Momenta

## Tensor Product of Lie-Group Representations

Recall from quantum mechanics that two kinematically independent systems with state spaces $\mathcal{H}_{1}$ and $\mathcal{H}_{2}$ can be looked at as one unified system with state space

$$
\mathcal{H}=\mathcal{H}_{1} \otimes \mathcal{H}_{2}
$$

An element $\left|\psi_{1} \psi_{2}\right\rangle \equiv\left|\psi_{1}\right\rangle \otimes\left|\psi_{2}\right\rangle \in \mathcal{H}$ is given by sums $\sum_{i} \psi_{1 i} \otimes \psi_{2 i}$.
The tensor product of two operators $\hat{\mathrm{A}}_{1}$ and $\hat{\mathrm{A}}_{2}$ on $\mathcal{H}_{1}$ and $\mathcal{H}_{2}$ is defined by

$$
\left(\hat{\mathrm{A}}_{1} \otimes \hat{\mathrm{~A}}_{2}\right)\left|\psi_{1}\right\rangle \otimes\left|\psi_{2}\right\rangle:=\hat{\mathrm{A}}_{1}\left|\psi_{1}\right\rangle \otimes \hat{\mathrm{A}}_{2}\left|\psi_{2}\right\rangle .
$$

If these operators are given by some irreducible representations $D^{1}, D^{2}$ of a group $G$

$$
\hat{\mathrm{A}}_{i}=\mathrm{i} \hbar \mathrm{D}^{i}(t),
$$

then a tensor product between Lie-group representations is defined by

$$
\left(\mathrm{D}^{1} \otimes \mathrm{D}^{2}\right)(t)\left|\psi_{1}\right\rangle \otimes\left|\psi_{2}\right\rangle:=\mathrm{D}^{1}(t)\left|\psi_{1}\right\rangle \otimes \mathrm{D}^{2}(t)\left|\psi_{2}\right\rangle
$$

## Clebsch-Gordan Series for $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$

A tensor product of irreducible representations can be expanded as a direct sum of irreducible representations, as mentioned in Corollary 2.10:

$$
\mathrm{D}^{j_{1}} \otimes \mathrm{D}^{j_{2}}=\bigoplus_{j} \mathrm{D}^{j}
$$

Our task is to determine exactly which irreducible representations $\mathrm{D}^{j}$ will sum up to give the tensor product $\mathrm{D}^{j_{1}} \otimes \mathrm{D}^{j_{2}}$. We will do this for $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$.

Theorem 5.1 The tensor product of two irreducible representations $\mathrm{D}^{j_{1}}$ and $\mathrm{D}^{j_{2}}$ of $\mathrm{SU}(2)$ and (if $j_{1}, j_{2} \in \mathbb{N}$ ), $\mathrm{SO}(3)$ has the following decomposition

$$
\mathrm{D}^{j_{1}} \otimes \mathrm{D}^{j_{2}}=\bigoplus_{j=\left|j_{1}-j_{2}\right|}^{j_{1}+j_{2}} \mathrm{D}^{j}
$$

This is called the Clebsch-Gordan series.

Proof: See Jones §6.2.
Examples:

$$
\begin{aligned}
& \mathrm{D}^{\frac{1}{2}} \otimes \mathrm{D}^{\frac{1}{2}}=\bigoplus_{j=0}^{1} \mathrm{D}^{j}=\mathrm{D}^{1} \oplus \mathrm{D}^{0} \\
& \mathrm{D}^{1} \otimes \mathrm{D}^{\frac{1}{2}}=\bigoplus_{j=\frac{1}{2}}^{\frac{3}{2}} \mathrm{D}^{j}=\mathrm{D}^{\frac{3}{2}} \oplus \mathrm{D}^{\frac{1}{2}} \\
& \mathrm{D}^{5} \otimes \mathrm{D}^{2}=\bigoplus_{j=3}^{7} \mathrm{D}^{j}=\mathrm{D}^{7} \oplus \mathrm{D}^{6} \oplus \mathrm{D}^{5} \oplus \mathrm{D}^{4} \oplus \mathrm{D}^{3}
\end{aligned}
$$

## Tensor Product of Lie-Algebra Representations

Let $L^{1}, L^{2}$ be irreducible representations of $\mathfrak{g}$, the Lie-algebra corresponding to $G$. We define a representation $L^{1} \otimes L^{2}$ of $\mathfrak{g}$ by

$$
\left(\mathrm{L}^{1} \otimes \mathrm{~L}^{2}\right)(X):=\mathrm{L}^{1}(X) \otimes \mathbb{1}+\mathbb{1} \otimes \mathrm{L}^{2}(X)
$$

This is designed so that when $L^{i}$ is the representation of $\mathfrak{g}$ induced by the representation $D^{i}$ of $G$ then $L^{1} \otimes L^{2}$ is induced by $D^{1} \otimes D^{2}$. Beware that the tensor product symbol $\otimes$ has different meanigs, depending on whether one is dealing with Lie groups or with Lie algebras.
For example, consider a system of two non-interacting particles with spins $s_{1}$ and $s_{2}$. The total spin operator $\hat{\mathbf{S}}$ is given by the representation D of $\operatorname{su}(2)$ in the following way:

$$
\hat{\mathbf{S}}=\hat{\mathbf{S}}_{1} \otimes \mathbb{1}+\mathbb{1} \otimes \hat{\mathbf{S}}_{2}=\mathrm{i} \hbar \bigoplus_{s=\left|s_{1}-s_{2}\right|}^{s_{1}+s_{2}} \mathrm{D}^{s}\left(\mathbf{S}_{\mathbf{n}}\right)
$$

For the special case of two spin- $\frac{1}{2}$ particles, the total spin operator is given by

$$
\hat{\mathbf{S}}=\mathrm{i} \hbar\left(\mathrm{D}^{0} \oplus \mathrm{D}^{1}\right)\left(\mathbf{s}_{\mathbf{n}}\right) .
$$

## Clebsch-Gordan Coefficients

Although we can "reduce" $\mathrm{D}^{j_{1}} \otimes \mathrm{D}^{j_{2}}$ into irreducible components $\bigoplus \mathrm{D}^{j}$, we still need to find an appropriate basis if we want to express these matrices into the block-diagonal form given in Theorem 2.2.
We start with the basis of eigenvectors

$$
\left|j_{1} m_{1} j_{2} m_{2}\right\rangle \equiv\left|j_{1} m_{1}\right\rangle \otimes\left|j_{2} m_{2}\right\rangle
$$

which span $\mathcal{H}$. The operators

$$
\left(\hat{\mathbf{J}}_{1}\right)^{2},\left(\hat{\mathbf{J}}_{2}\right)^{2},(\hat{\mathbf{J}})^{2},(\hat{\mathrm{~J}})_{3}
$$

form a complete system of commuting observables. Therefore, $\mathcal{H}$ can be spanned by a basis of common eigenvectors

$$
\left|j_{1} j_{2} j m\right\rangle
$$

In order to determine the new basis $\left\{\left|j_{1} j_{2} j m\right\rangle\right\}$, we express it in terms of the old basis $\left\{\left|j_{1} m_{1} j_{2} m_{2}\right\rangle\right\}$ :
Theorem 5.2 The two bases $\left\{\left|j_{1} m_{1} j_{2} m_{2}\right\rangle\right\}$ and $\left\{\left|j_{1} j_{2} j m\right\rangle\right\}$ are related by

$$
\left|j_{1} j_{2} j m\right\rangle=\sum_{m_{1}=-j_{1}}^{j_{1}} \sum_{m_{2}=-j_{2}}^{j_{2}} c\left(j_{1}, j_{2}, j, m_{1}, m_{2}, m\right)\left|j_{1} m_{1} j_{2} m_{2}\right\rangle
$$

where $c\left(j_{1}, j_{2}, j, m_{1}, m_{2}, m\right) \in \mathbb{R}$ are called the Clebsch-Gordan coefficients.
These coefficients can either be calculated for fixed $j_{1}$ and $j_{2}$ as in the following example, or looked up in a table.

## Example: Two Particles of spin $\frac{1}{2}$

The components of the total spin operator are given by

$$
(\hat{\mathbf{S}})_{i}=\mathrm{i} \hbar\left(\mathrm{D}^{\frac{1}{2}} \otimes \mathrm{D}^{\frac{1}{2}}\right)\left(\mathbf{s}_{i}\right)=\mathrm{i} \hbar\left(\mathrm{D}^{1} \oplus \mathrm{D}^{0}\right)\left(\mathbf{s}_{i}\right) .
$$


$D^{\frac{1}{2}}$
$D^{\frac{1}{2}}$
$D^{1} \quad D^{0}$

Figure 5.1: The representations $D^{\frac{1}{2}} \otimes D^{\frac{1}{2}}$ and $D^{1} \oplus D^{0}$ on the four-dimensional spinor space

Given the old basis (we simplify notation by omitting $j_{1}$ and $j_{2}$ on the r.h.s.)

$$
\begin{aligned}
\left|\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right\rangle & \equiv|++\rangle \\
\left|\frac{1}{2}, \frac{1}{2}, \frac{1}{2},-\frac{1}{2}\right\rangle & \equiv|+-\rangle \\
\left|\frac{1}{2},-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right\rangle & \equiv|-+\rangle \\
\left|\frac{1}{2},-\frac{1}{2}, \frac{1}{2},-\frac{1}{2}\right\rangle & \equiv|--\rangle,
\end{aligned}
$$

we'd like to find a new basis

$$
\begin{aligned}
\left|\frac{1}{2}, \frac{1}{2}, 1,1\right\rangle & \equiv|11\rangle \\
\left|\frac{1}{2}, \frac{1}{2}, 1,0\right\rangle & \equiv|10\rangle \\
\left|\frac{1}{2}, \frac{1}{2}, 1,-1\right\rangle & \equiv|1-1\rangle \\
\left|\frac{1}{2}, \frac{1}{2}, 0,0\right\rangle & \equiv|00\rangle .
\end{aligned}
$$

We start with $|++\rangle$ which is the tensor product of the highest-weight vectors $\left|\frac{1}{2}, \frac{1}{2}\right\rangle$. One can check that

$$
\hat{\mathbf{J}}^{2}|++\rangle=2 \hbar^{2}|++\rangle, \quad \hat{\mathrm{J}}_{3}|++\rangle=\hbar|++\rangle .
$$

Hence, we can define

$$
|11\rangle:=|++\rangle
$$

and climb down using the ladder operator $\hat{\mathrm{S}}_{-}$. Application of $\hat{\mathrm{S}}_{-}$to $|11\rangle$ and using the formula given in Theorem 4.1 yields

$$
\hat{\mathrm{S}}_{-}|11\rangle=\sqrt{2}|10\rangle, \quad \hat{\mathrm{S}}_{-}^{2}|11\rangle=2|1-1\rangle .
$$

Application to $|++\rangle$ yields
and

$$
\begin{aligned}
\hat{\mathrm{S}}_{-}^{2}|++\rangle= & \hat{\mathrm{S}}_{-}(|-+\rangle+|+-\rangle) \\
= & \left(\hat{\mathrm{S}}_{1-}\left|\frac{1}{2},-\frac{1}{2}\right\rangle\right) \otimes\left|\frac{1}{2}, \frac{1}{2}\right\rangle+\left|\frac{1}{2},-\frac{1}{2}\right\rangle \otimes\left(\hat{\mathrm{S}}_{2-}\left|\frac{1}{2}, \frac{1}{2}\right\rangle\right) \\
& +\left(\hat{\mathrm{S}}_{1-}\left|\frac{1}{2}, \frac{1}{2}\right\rangle\right) \otimes\left|\frac{1}{2},-\frac{1}{2}\right\rangle+\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes\left(\hat{\mathrm{S}}_{2-}\left|\frac{1}{2},-\frac{1}{2}\right\rangle\right) \\
= & 0+\left|\frac{1}{2},-\frac{1}{2}\right\rangle \otimes\left|\frac{1}{2},-\frac{1}{2}\right\rangle+\left|\frac{1}{2},-\frac{1}{2}\right\rangle \otimes\left|\frac{1}{2},-\frac{1}{2}\right\rangle+0 \\
= & 2|--\rangle .
\end{aligned}
$$

Thus, we read off

$$
|10\rangle=\frac{1}{\sqrt{2}}(|-+\rangle+|+-\rangle), \quad|1-1\rangle=|--\rangle
$$

As for $|00\rangle$, we know that it must be a linear combination of the old basis:

$$
|00\rangle=A|++\rangle+B|+-\rangle+C|-+\rangle+D|--\rangle,
$$

where $|A|^{2}+|B|^{2}+|C|^{2}+|D|^{2}=1$. As the eigenspaces of the self-adjoint operator $\hat{\mathbf{J}}^{2}$ are orthogonal, $|00\rangle$ must be orthogonal to $|11\rangle,|10\rangle$ and $|1-1\rangle$. This yields

$$
\langle 11 \mid 00\rangle=A=0, \quad\langle 10 \mid 00\rangle=B+C=0, \quad\langle 1-1 \mid 00\rangle=D=0 .
$$

In accordance with the Condon-Shortley convention (c.f. Cornwell, §12.5) we choose $B$ to be positive. Then normalization yields

$$
|00\rangle=\frac{1}{\sqrt{2}}(|+-\rangle-|-+\rangle) .
$$

Thus, we have fully expressed the new basis $\left|j_{1} j_{2} j m\right\rangle$ in terms of the old basis $\left|j_{1} m_{1} j_{2} m_{2}\right\rangle$ :

Notice that the coefficients of these linear combinations match up with the values found in the table of Clebsch-Gordan coefficients for $\frac{1}{2} \otimes \frac{1}{2}$. See Table 12.1 of Cornwell for some simple tables of Clebsch-Gordan coefficients.

## Chapter 6

## The Hidden Symmetry of the Hydrogen Atom

In this chapter, we briefly consider the hydrogen atom. Recall that a rotationally symmetric potential $V(r)$ allows us to split up $L^{2}\left(\mathbb{R}^{3}\right), \mathrm{d} \mu$ into radial and angular components $\psi=f \otimes \phi$ and that the radial part can be solved by the Hamilton operator

$$
\hat{\mathrm{H}} f=E f
$$

Now consider the special case of the Coulomb potential for the hydrogen atom

$$
V(r)=-\frac{e^{2}}{r} \quad \hat{\mathrm{H}}=\frac{\hat{\mathbf{p}}^{2}}{2 \mu}-\frac{e^{2}}{r} .
$$

This potential is said to possess some "hidden" symmetry besides the obvious rotational invariance. We will explain this for bound states. Let $E<0$ be fixed and consider the operator

$$
\hat{\mathbf{B}}=\frac{1}{2 e^{2} \mu}(\hat{\mathbf{L}} \times \hat{\mathbf{p}}-\hat{\mathbf{p}} \times \hat{\mathbf{L}})+\frac{\hat{\mathbf{r}}}{r},
$$

acting on the angular part $\phi$. This is the quantum-mechanical version of the Lenz-Runge vector, which is a conserved quantity for a $\frac{1}{r}$ potential. Define the reduced Lenz-Runge operator by

$$
\hat{\mathbf{A}}=\sqrt{-\frac{\mu e^{4}}{2 E}} \hat{\mathbf{B}} \equiv \sqrt{-\frac{\gamma}{E}} \hat{\mathbf{B}}
$$

where $\gamma \equiv \frac{\mu e^{4}}{2}$. We then have the following commutation relations:

$$
\begin{aligned}
{\left[\hat{\mathrm{L}}_{i}, \hat{\mathrm{~L}}_{j}\right] } & =\mathrm{i} \hbar \epsilon_{i j k} \hat{\mathrm{~L}}_{k} \\
{\left[\hat{\mathrm{~A}}_{i}, \hat{\mathrm{~A}}_{j}\right] } & =\mathrm{i} \hbar \epsilon_{i j k} \hat{\mathrm{~L}}_{k} \\
{\left[\hat{\mathrm{~L}}_{i}, \hat{\mathrm{~A}}_{j}\right] } & =\mathrm{i} \hbar \epsilon_{i j k} \hat{\mathrm{~A}}_{k}
\end{aligned}
$$

The second equation shows us that the vector operator $\hat{\mathbf{A}}$ fulfills the condition of a general angular momentum operator, and thus the vector operator

$$
\hat{\mathbf{Y}}=\frac{1}{\mathrm{i} \hbar} \hat{\mathbf{A}}
$$

generates an su(2) Lie algebra

$$
\left[\hat{\mathrm{Y}}_{i}, \hat{Y}_{j}\right]=\epsilon_{i j k} \hat{\mathrm{Y}}_{k}
$$

in the same way that $\hat{\mathbf{X}}=\frac{1}{\mathrm{i} \hbar} \hat{\mathbf{L}}$ does. We now consider the operators

$$
\hat{\mathbf{M}} \equiv \frac{1}{2}(\hat{\mathbf{L}}+\hat{\mathbf{A}}), \quad \hat{\mathbf{N}}=\frac{1}{2}(\hat{\mathbf{L}}-\hat{\mathbf{A}}) .
$$

They fulfill the commutation relations

$$
\begin{aligned}
{\left[\hat{\mathrm{M}}_{i}, \hat{\mathrm{M}}_{j}\right] } & =\mathrm{i} \hbar \epsilon_{i j k} \hat{\mathrm{M}}_{k} \\
{\left[\hat{\mathrm{~N}}_{i}, \hat{\mathrm{~N}}_{j}\right] } & =\mathrm{i} \hbar \epsilon_{i j k} \hat{\mathrm{~N}}_{k} \\
{\left[\hat{\mathrm{M}}_{i}, \hat{\mathrm{~N}}_{j}\right] } & =0
\end{aligned}
$$

Wee see that $\hat{\mathbf{M}}$ and $\hat{\mathbf{N}}$ fulfill the condition of a general angular momentum operatoreven though their physical meaning is not evident. And because they commute (unlike $\hat{\mathbf{A}}$ and $\hat{\mathbf{L}}$ ), they form a direct sum of two independent su(2) algebras

$$
\mathrm{su}(2) \oplus \mathrm{su}(2)
$$

Theorem 6.1 We have

$$
\operatorname{su}(2) \oplus \operatorname{su}(2) \cong \mathrm{so}(4)
$$

Proof: Cornwell, Appendix G, Section 2b
Thus, the Lie algebra of the hydrogen atom is so(4), whose corresponding Lie group is $\mathrm{SO}(4)$, the group of proper rotations in four dimensions. While the Lie group $\mathrm{SO}(3)$ expresses the symmetry of a system with a rotationally invariant potential $V(r)$, the special case of a $\frac{1}{r}$-potential provides an extra "accidental" symmetry, which can only be explained by considering the Lie group $\mathrm{SO}(4)$. Notice that we have arrived at the group SO(4) indirectly, that is, through its Lie algebra so(4). A more direct approach without taking this Lie-algebra "shortcut" would be quite difficult to grasp-imagine having to visualize a four-dimensional rotation!

## The Quantum Number $n$

The operators $\hat{\mathrm{H}}, \hat{\mathrm{M}}^{2}, \hat{\mathrm{M}}_{3}, \hat{\mathbf{N}}^{2}, \hat{\mathrm{~N}}_{3}$ form a system of commuting ovesrvables, and we have eigenvectors

$$
\left|E \alpha m_{\alpha} \beta m_{\beta}\right\rangle \equiv|E\rangle
$$

Notice that this is an irreducible basis for a $(2 \alpha+1)(2 \beta+1)$-dimensional representation of so(4). Also notice that $\alpha, \beta=0, \frac{1}{2}, 1, \ldots$ act as a kind of spin number as in Theorem 4.1. The Casimir operators give us

$$
\begin{aligned}
\hat{\mathbf{M}}^{2}|E\rangle & =\hbar^{2} \alpha(\alpha+1)|E\rangle \\
\hat{\mathbf{N}}^{2}|E\rangle & =\hbar^{2} \beta(\beta+1)|E\rangle
\end{aligned}
$$

Using the identity

$$
\hat{\mathbf{A}} \cdot \hat{\mathbf{L}}=0,
$$

we have

$$
\hat{\mathbf{M}}^{2}=\hat{\mathbf{N}}^{2}=\frac{1}{4}\left(\hat{\mathbf{A}}^{2}+\hat{\mathbf{L}}^{2}\right)
$$

which, when applied to $|E\rangle$, gives us

$$
\alpha=\beta
$$

Using the identity

$$
\begin{aligned}
& \hat{\mathbf{B}}^{2}=\mathbb{1}+\frac{2}{\mu e^{4}} \hat{\mathrm{H}}\left(\hat{\mathbf{L}}^{2}+\hbar^{2} \mathbb{I}\right), \\
& \hat{\mathbf{A}}^{2}=-\frac{\gamma}{E} \mathbb{1}-\frac{\hat{\mathrm{H}}}{E}\left(\hat{\mathbf{L}}^{2}+\hbar^{2} \mathbb{I}\right)
\end{aligned}
$$

and $\left[\hat{\mathrm{H}}, \hat{\mathbf{L}}^{2}\right]=0$, we obtain

$$
4 \hat{\mathbf{M}}^{2}=\hat{\mathbf{A}}^{2}+\hat{\mathbf{L}}^{2}=-\frac{\gamma}{E} \mathbb{1}-\frac{\hat{\mathrm{H}}}{E} \hat{\mathbf{L}}^{2}-\hbar^{2} \frac{\hat{\mathrm{H}}}{E}+\hat{\mathbf{L}}^{2}
$$

Applying this to $|E\rangle$ gives

$$
\begin{aligned}
4 \hbar^{2} \alpha(\alpha+1) & =-\hbar^{2}-\frac{\gamma}{E}=-\hbar^{2}-\frac{\mu e^{4}}{2 E} \\
E & =-\frac{\mu e^{4}}{2 \hbar^{2}(2 \alpha+1)^{2}}
\end{aligned}
$$

Defining $n \equiv 2 \alpha+1=1,2,3, \ldots$ gives us the well-known result for the energy levels of the hydrogen atom:

$$
E_{n}=-\frac{\mu e^{4}}{2 \hbar^{2} n^{2}}
$$

## The Quantum Number $l$

The eigenvectors of fixed energy $E_{n}$ form a basis for a $(2 \alpha+1)(2 \beta+1)=(2 \alpha+1)^{2}=n^{2}$ dimensional irreducible representation of so(4). Notice that this corresponds to the $n^{2}$ degeneracy of the energy levels of the hydrogen atom.
Going back to the defining equations of $\hat{\mathbf{M}}$ and $\hat{\mathbf{N}}$, we see that

$$
\hat{\mathbf{L}}=\hat{\mathbf{M}}+\hat{\mathbf{N}}
$$

or in other words, $\hat{\mathbf{L}}$ is the sum of independent angular momenta, which corresponds to the tensor product of representations $\mathrm{D}^{\alpha}$ and $\mathrm{D}^{\beta}=\mathrm{D}^{\alpha}$ :

$$
\mathrm{D}^{\alpha} \otimes \mathrm{D}^{\alpha}=\bigoplus_{l=0}^{2 \alpha} D^{l}=\mathrm{D}^{n-1} \oplus \cdots \oplus \mathrm{D}^{0}
$$

That is, the quantum number $l$ takes on the well-known values

$$
l=n-1, n-2, \ldots, 1,0 .
$$

## Glossary

| abelian | abelsch, 2 |
| :---: | :---: |
| algebra | Algebra, 16 |
| algebra multiplication | Algebramultiplikation, 16 |
| associative algebra | assoziative Algebra, 16 |
| bijective | bijektiv, 14 |
| carrier space | Trägerraum, 21 |
| Casimir operator | Casimir-Operator, 46 |
| character | Charakter, 24 |
| Clebsch-Gordan coefficients | Clebsch-Gordan Koeffizienten, 61 |
| closure | Abgeschlossenheit, 2 vollständig reduzibel, 22 |
| complex matrix representation | komplexe Matrixdarstellung, 24 |
| conjugacy class | Konjugationsklasse, 8 |
| conjugate | konjugiert, 7 |
| connected | zusammenhängend, 33 |
| continuous | kontinuierliche, 2 |
| coset | Nebenklasse, 8 |
| countable | abzählbare, 2 |
| cyclic group | zyklische Gruppe, 3 |
| dihedral group | Diedergruppe, 5 |
| dimension | Dimension, 24 |
| equivalence relation equivalent | Äquivalenzrelation, 7 äquivalent, 24 |
| factor group | Faktorgruppe, 9 |
| faithful | treu, 21 |
| general linear group | allgemeine lineare Gruppe, 19 |
| generator | Generator, 3 |
| group | Gruppe, 1 |
| group multiplication | Gruppenmultiplikation, 1 |


| highest-weight vector | Vektor des höchsten Gewichts, 48 |
| :--- | :--- |
| homomorphism | Homomorphismus, 12 |
| identity | Einheitselement, 1 |
| image | Bild, 13 |
| infinitesimal generator | infinitesimale Generator, 33 |
| injective | injektiv, 13 |
| invariant | invariant, 22 |
| inverse | inverses Element, 1 |
| invertible | invertibel, 14 |
| irreducible | irreduzibel, 22 |
| isomorphic |  |
| isomorphism | isomorph, 14 |
| kernel | Isomorphismus, 14 |
| ladder operators | Kern, 13 |
| Latin square <br> left coset <br> Lie algebra <br> Lie bracket <br> Lie group <br> lowest-weight vector | Leiteroperatoren, 47 |
| multiplication tables | Linksnebenklasse, 9 |


| simply connected | einfach zusammenhängend, 33 |
| :--- | :--- |
| spherical harmonics | Kugelfächenfunktionen, 57 |
| spin number | Spinzahl, 50 |
| spinor representation | Spinordarstellung, 44 |
| spinor space | Spinorraum, 50 |
| structure constants | Strukturkonstanten, 17 |
| subgroup | Untergruppe, 7 |
| surjective | surjektiv, 14 <br> symplectic group |
| symplektische Gruppe, 19 |  |
| tangent space <br> trivial subgroups <br> two-fold covering | Tangentialraum, 34 |
| triviale Untergruppen, 7 |  |
| unitary | zwei-blättrige Überlagerung, 43 |
| unitary group | unitär, 21 |
| unitäre Gruppe, 19 |  |
| vector addition | Vektoraddition, 16 |
| vector space | Vektorraum, 16 |

