

Chapter 2

A Little Bit about Groups

We shall frequently use the language of groups. Not only are group theoretical concepts important in general relativity; the ideas pervade modern physics, from crystallography to particle physics. Therefore, we shall take out a little time now in order to introduce the concept of a group and some of the mathematical techniques associated with group theory.

We shall begin with a formal definition such as one might find in a mathematics book.

Def. 1 *A group consists of a set G together with a binary operation, called group multiplication, and usually denoted by juxtaposition, such that the following properties hold:*

- (1) **Associative Law:** $(ab)c = a(bc)$ for all $a, b, c \in G$,
- (2) **Existence of Identity Element:** *There exists an element $e \in G$ such that, for all $a \in G$, $ea = a$,*
- (3) **Existence of Inverse:** *For each $a \in G$, there exists an element $b \in G$ such that $ba = e$. (This inverse element b is usually denoted by the symbol a^{-1} .)*

For a given operation of group multiplication we shall often denote the group by the same symbol G used to denote the set of elements of the group. In general group multiplication is noncommutative; i.e., $ab \neq ba$. A group that is commutative is called an *Abelian group* after the Norwegian mathematician Niels Hendrik Abel (1802-1829).¹

Note that in stating property (2) we did not stipulate that $ae = a$. Nor did we state that the element e is unique. Moreover, in stating property (3)

¹Some people make the most of a short life!

we did not stipulate that $ab = e$. Nor did we state that the element b is unique. You will find the following to be a challenging exercise.

Ex. 1 *If G is a group, prove that*

- (a) *for all $a \in G$, $ae = a$.*
- (b) *for each $a \in G$, $aa^{-1} = e$.*
- (c) *the element e is unique.*
- (d) *for each $a \in G$, the element a^{-1} is unique.*

If the exercise appears to be simple, make certain you have not used in your “proof” one of the properties of groups that was to be established. If you are unable to concoct a satisfactory proof, you can find such a proof on the first page of the book *Topological Groups* by Pontriagin (Princeton University Press).

Spatial translations, rotations about a fixed axis, and Galilean transformations are examples of Abelian groups. An example of a group which is not Abelian is the rotation group in three dimensions, usually designated by $O(3)$ because it is also called the *orthogonal group*.

To see that $O(3)$ is indeed a noncommutative group, place a book on a table in front of you so that you are looking down at the cover. Rotate the book ninety degrees counterclockwise about the vertical axis. The binding of the book should now be toward your stomach. Next, rotate the book ninety degrees away from you about a left-right axis. You should now be staring down at the binding of the book.

Now, return the book to its original orientation, so you are again staring down at the cover. Carry out the two rotations in the reverse order. The rotation of the book away from you about a left-right axis results in your looking down at the bottom edge of the book, and the following counter clockwise rotation about the vertical axis leaves you looking down at the bottom edge of the book, *not* at the binding of the book.

The groups we have been talking about thus far are groups with infinitely many elements. To acquire more rapidly an understanding of what a group is, let us consider some very simple groups that have a finite number of elements.

If a group consists of only one element that single element must be the identity element e . If a group has two distinct elements e, a , where $a \neq e$, then necessarily $a^2 = e$. This group is Abelian. The only group with three distinct elements consists of the elements e, a, a^2 , where $a^3 = e$. This group is also Abelian. In fact, if we consider any group with n elements, where n is prime, that group must be the so-called *cyclic group* with elements $e, a, a^2, \dots, a^{n-1}$, where $a^n = e$.

The simplest noncommutative group has six distinct elements and the following multiplication table:

	e	a	b	c	d	f
e	e	a	b	c	d	f
a	a	b	e	d	f	c
b	b	e	a	f	c	d
c	c	f	d	e	b	a
d	d	c	f	a	e	b
f	f	d	c	b	a	e

Ex. 2 Show that the set of elements $\{e, a, b, c, d, f\}$ with the above multiplication table constitutes a group.

Ex. 3 Think about the symmetries of an equilateral triangle, trying to find some relationship of these symmetries to the group studied in the preceding exercise.

Ex. 4 Find two distinct noncommutative groups with eight elements; i.e., write out the multiplication tables for the two groups. (Check carefully that the two groups are really distinct and that you have not simply named the elements of the same group in two different ways.)

When one turns one's attention from finite groups to continuous groups; i.e., groups with infinitely many elements, it becomes impractical to give such group multiplication tables. One must find other ways to describe the structure of such groups.

The two-dimensional rotation group, designated $O(2)$, is an Abelian group. It can be described in terms of transformations upon a reference system (frame) consisting of two orthogonal unit vectors, \mathbf{e}_1 and \mathbf{e}_2 . Under a rotation through angle θ ,

$$\begin{aligned}\mathbf{e}_1 &\rightarrow \mathbf{e}'_1 = \mathbf{e}_1 \cos \theta + \mathbf{e}_2 \sin \theta \\ \mathbf{e}_2 &\rightarrow \mathbf{e}'_2 = -\mathbf{e}_1 \sin \theta + \mathbf{e}_2 \cos \theta.\end{aligned}$$

Each value of θ corresponds to a different element of the group $O(2)$. Under a second rotation through angle θ' ,

$$\begin{aligned}\mathbf{e}'_1 \rightarrow \mathbf{e}''_1 &= \mathbf{e}'_1 \cos \theta' + \mathbf{e}'_2 \sin \theta' \\ \mathbf{e}'_2 \rightarrow \mathbf{e}''_2 &= -\mathbf{e}'_1 \sin \theta' + \mathbf{e}'_2 \cos \theta'.\end{aligned}$$

These two transformations, when performed in succession, give

$$\begin{aligned}\mathbf{e}_1 \rightarrow \mathbf{e}''_1 &= (\mathbf{e}_1 \cos \theta + \mathbf{e}_2 \sin \theta) \cos \theta' + (-\mathbf{e}_1 \sin \theta + \mathbf{e}_2 \cos \theta) \sin \theta' \\ &= \mathbf{e}_1 \cos(\theta + \theta') + \mathbf{e}_2 \sin(\theta + \theta'), \\ \mathbf{e}_2 \rightarrow \mathbf{e}''_2 &= -(\mathbf{e}_1 \cos \theta + \mathbf{e}_2 \sin \theta) \sin \theta' + (-\mathbf{e}_1 \sin \theta + \mathbf{e}_2 \cos \theta) \cos \theta' \\ &= -\mathbf{e}_1 \sin(\theta + \theta') + \mathbf{e}_2 \cos(\theta + \theta').\end{aligned}$$

This, however, is equivalent to a single rotation through angle $\theta + \theta'$. We are dealing here with an Abelian group, the identity element of which corresponds to $\theta = 0$. The inverse element to the element corresponding to θ is the element corresponding to $-\theta$.

The *abstract group* consists of the rotations themselves, while the transformations induced upon the basic vectors $\mathbf{e}_1, \mathbf{e}_2$ constitute a particular *realization* (or *representation*) of the abstract group. If one had used a different orthonormal basis, or one had used some non-orthonormal basis, one would have constructed a different realization of the same abstract group, $O(2)$.

The angle θ was used here as a *parameter* to label an element of the group. One could, of course, parametrize the group using, say, θ^3 in place of θ . An advantage to using θ is that group multiplication corresponds to simple addition of the θ -parameters.

The three-dimensional rotation group, designated $O(3)$, is a group that is non-commutative. It can be described in terms of transformations upon a reference system (frame) of three orthonormal unit vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. However, the transformation formulas are in this case far from obvious, for even the natural choice of parameters is not very obvious. One possible choice consists of the three Euler angles, familiar from the treatment of spinning tops in classical mechanics. Instead of using Euler angles, we shall explore a more systematic approach, for our ultimate goal is to understand the nature of the Lorentz group, not just $O(3)$.

What we are looking for here are 3×3 matrices R_{ab} such that the transformations

$$\mathbf{e}_a \rightarrow \mathbf{e}'_a = \sum_b R_{ab} \mathbf{e}_b$$

constitute a realization of $O(3)$. In the case of $O(2)$ the matrices were given by

$$R_{ab} \rightarrow \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \quad (2.1)$$

How might these matrices have been constructed systematically, if their form had not been so obvious?

The key point is that the matrices R_{ab} of any $O(n)$ group are *orthogonal* matrices, i.e.,

$$\sum_b R_{ba} R_{bc} = \delta_{ac}.$$

This follows from the fact that rotations leave all scalar products invariant. In terms of matrix language, $R^T R = I$, where R^T is the transpose of R and where I denotes the unit matrix. If one further restricts R so that $\det R = +1$, then the group is designated by $SO(n)$ and is called the *special* orthogonal group.

The trick which we shall invoke in order to construct the matrices R is to express these matrices as *formal* exponentials. We shall write

$$R = e^{-iX} := I - iX + \frac{1}{2}(-iX)^2 + \dots,$$

where the formal symbol e^{-iX} is defined by the power series expansion, the sum of terms each one of which is an $n \times n$ matrix.

The reality of the matrices R implies that the matrices X are pure imaginary, while the orthogonality of R implies that the matrix X is skew-symmetric. In two dimensions, then, X must have the form

$$X \rightarrow \begin{pmatrix} 0 & i\theta \\ -i\theta & 0 \end{pmatrix},$$

where θ is a real parameter (which will turn out to be the angle of rotation). The matrix R is constructed by summing the series that defines e^{-iX} . The result, which is very easy to obtain, is just what we expected all along.

Ex. 5 Using the power series expansions of $\cos \theta$ and $\sin \theta$, verify that the 2×2 matrix R has the expected form (2.1).

If we now turn our attention to $SO(3)$, we must deal with 3×3 matrices X which are pure imaginary and skew-symmetrical. Obviously, *three* parameters are necessary. We can write, for example,

$$X = \theta \mathbf{n} \cdot \mathbf{S},$$

where θ is a real parameter and \mathbf{n} is a real unit vector (corresponding to two more real parameters). The matrices S_1, S_2 and S_3 are defined by

$$\begin{aligned} S_1 &\rightarrow \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \\ S_2 &\rightarrow \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \\ S_3 &\rightarrow \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

In principle one can now sum the series and thus construct the 3×3 matrices R . However, if you attempt it, you will find that it is not quite as trivial as it was in the 2×2 case. How then can one obtain the matrices R appropriate for $SO(3)$?

Let us digress and consider the commutation relations which are satisfied by the matrices S_1, S_2 and S_3 . It is not difficult to verify that

$$\begin{aligned} S_2 S_3 - S_3 S_2 &= i S_1, \\ S_3 S_1 - S_1 S_3 &= i S_2, \\ S_1 S_2 - S_2 S_1 &= i S_3. \end{aligned}$$

You may remember these relations as the *angular momentum commutation relations* of elementary quantum mechanics (with \hbar replaced by 1). Many different sets of three matrices satisfy these commutation relations. All such sets correspond to *realizations* (or *representations*) of one abstract *Lie algebra*. The matrices of lowest dimensionality that satisfy these relations are the 2×2 Pauli spin matrices (times $1/2$). These Pauli matrices are usually written as follows:

$$\begin{aligned} \sigma_1 &\rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_2 &\rightarrow \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \sigma_3 &\rightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \end{aligned}$$

and they also satisfy the simple relations $\sigma_2 \sigma_3 = i \sigma_1, \sigma_3 \sigma_1 = i \sigma_2, \sigma_1 \sigma_2 = i \sigma_3$. The square of each Pauli matrix is the unit matrix, and any two different Pauli matrices anticommute.

The 2×2 matrices $A = e^{-i\theta \mathbf{n} \cdot (\frac{1}{2}\boldsymbol{\sigma})}$ are readily constructed, unlike the 3×3 matrices $R = e^{-i\theta \mathbf{n} \cdot \mathbf{S}}$. The result, obtained by summing a series that defines a formal exponential, is simply

$$A = \cos \frac{\theta}{2} I - i \sin \frac{\theta}{2} \mathbf{n} \cdot \boldsymbol{\sigma}.$$

These matrices A constitute a group with the same abstract Lie algebra as the three dimensional rotation group. This group of complex 2×2 matrices is called $SU(2)$, where the U designates the unitary character of the matrices and the S signifies that $\det A = +1$. (Special unitary group!)

There exists a product-preserving correspondence between elements of $SO(3)$ and elements of $SU(2)$. When such a correspondence is one-to-one, it is called an *isomorphism*. In the present case we are dealing with a two-to-one correspondence, and as such it is called a *homomorphism*, not to be confused with the word homeomorphism, which is something completely different. We shall see that to each element of $SO(3)$ there correspond two elements of $SU(2)$, which just differ in overall sign. The existence of this homomorphism will permit us to construct the matrices of $SO(3)$ without difficulty!

Consider the 2×2 Hermitian matrix

$$Y \rightarrow \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}.$$

Because Y is traceless, it can be expanded in terms of the traceless Hermitian matrices σ_1, σ_2 and σ_3 . It is easy to see that

$$Y = x\sigma_1 + y\sigma_2 + z\sigma_3.$$

Note also that the trace of Y^2 is given by

$$\text{Tr}(Y^2) = 2(x^2 + y^2 + z^2).$$

A unitary transformation has the form

$$Y \rightarrow Y' = AY A^\dagger,$$

where $AA^\dagger = I$. The dagger denotes the Hermitian conjugate of a complex matrix, i.e., the transpose of the complex conjugate. Under a unitary transformation a Hermitian matrix Y is transformed into another Hermitian matrix Y' with the same trace (zero). Furthermore, the trace of Y'^2 remains the same as the trace of Y^2 . Thus, $x^2 + y^2 + z^2$ is an invariant.

Clearly we are very close to finding the precise relation between the $SU(2)$ matrices A and the rotation matrices R , which also leave $x^2 + y^2 + z^2$ invariant, when x, y and z are identified as the Cartesian components of a vector, $x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$.

If $Y' = x'\sigma_1 + y'\sigma_2 + z'\sigma_3$, then it follows that $x' = \frac{1}{2}\text{Tr}(\sigma_1 Y')$, etc., because $\text{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}$. Thus, for example,

$$\begin{aligned} x' &= \frac{1}{2}\text{Tr}(\sigma_1 A Y A^\dagger) \\ &= \frac{1}{2}\text{Tr}[\sigma_1 A(x\sigma_1 + y\sigma_2 + z\sigma_3)A^\dagger] \\ &= \frac{1}{2}\text{Tr}(\sigma_1 A \sigma_1 A^\dagger)x + \frac{1}{2}\text{Tr}(\sigma_1 A \sigma_2 A^\dagger)y + \frac{1}{2}\text{Tr}(\sigma_1 A \sigma_3 A^\dagger)z. \end{aligned}$$

One obtains analogous expressions for y' and z' . Recalling the definition of the matrix R , we conclude that

$$R_{ab} = \frac{1}{2}\text{Tr}(\sigma_a A^\dagger \sigma_b A).$$

This formula displays explicitly the 2 to 1 homomorphism between $SO(3)$ and $SU(2)$. Note that each 3×3 matrix R corresponds to two 2×2 matrices A and $-A$. Furthermore, plugging in the known form of the general $SU(2)$ matrix A , we can evaluate the trace and obtain the form of the general $SO(3)$ matrix R .

Ex. 6 Construct the matrix R explicitly in terms of θ and \mathbf{n} .

Ex. 7 Suppose the parameters (θ, \mathbf{n}) and (θ', \mathbf{n}') of two successive rotations are given. Find the parameters (θ'', \mathbf{n}'') of a single rotation that accomplishes the same thing. [HINT: Take advantage of the homomorphism.]

Having concluded these preliminaries, which I hope you will find useful not only in connection with relativity theory but also in connection with other aspects of physics, we can now turn our attention to the Lorentz group, which is really what we have been interested in right along. Remember, the Lorentz group is the group of transformations under which the Maxwell theory of electromagnetism remains invariant in form. This is the group that Einstein in 1905 associated with changes of inertial reference frames, abandoning the older Galilean group under which Newton's theory was invariant in form.