1.1 Symmetry and Groups

Symmetry and Transformation

Symmetry plays a central role in modern theoretical physics.[1]

As the etymologist tells us, symmetry ("equal measure") originates in geometry ("earth measure"). We have a sense that an isosceles triangle is more symmetrical than an arbitrary triangle and that an equilateral triangle is more symmetrical than an isosceles triangle. Going further, we feel that a pentagon is more symmetrical than a square, a hexagon more symmetrical than a pentagon, and an $(n+1)$-sided regular polygon is more symmetrical than an n-sided regular polygon. And finally, a circle is more symmetrical than any regular polygon.

The n-sided regular polygon is left unchanged by rotations through any angle that is an integer multiple of $2\pi/n$, and there are n of these rotations. The larger n is, the more such rotations there are. This is why mathematicians and physicists feel that the hexagon is more symmetrical than a pentagon: $6 > 5$, QED.

To quantify this intuitive feeling, we should thus look at the set of transformations that leave the geometrical figure unchanged (that is, invariant). For example, we can reflect the isosceles triangle across the median that divides it into equal parts (see figure 1a).

Call the reflection r; then the set of transformations that leave the isosceles triangle invariant is given by $\{I, r\}$, where I denotes the identity transformation, that is, the transformation that does nothing. A reflection followed by a reflection has the same effect as the identity transformation. We write this statement as $r \cdot r = I$.

In contrast, the equilateral triangle is left invariant not only by reflection across any of its three medians (figure 1b) but also by rotation R_1 through $2\pi/3 = 120^\circ$ around its center, as well as rotation R_2 through $4\pi/3 = 240^\circ$. The set of transformations that leave the equilateral triangle invariant is thus given by $\{I, r_1, r_2, r_3, R_1, R_2\}$. That this set is larger than the set in the preceding paragraph quantifies the feeling that the equilateral triangle is more symmetrical than the isosceles triangle. Note that $R_1 \cdot R_1 = R_2$ and that $R_1 \cdot R_2 = I$.

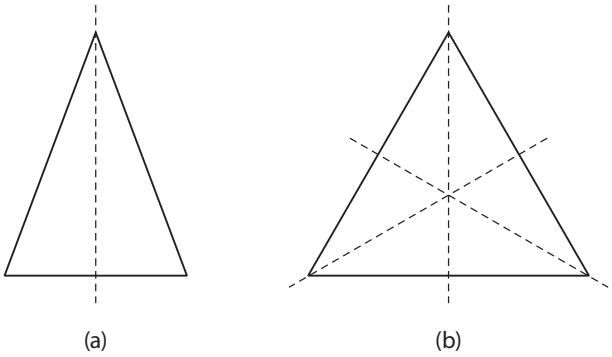

Figure 1

A circle is left invariant by an infinite number of transformations, namely, rotation $R(\theta)$ through any angle θ, and reflection across any straight line through its center. The circle is more symmetrical than any regular polygon, such as the equilateral triangle.

Symmetry in physics

In physics we are interested in the symmetries enjoyed by a given physical system. On a more abstract level, we are interested in the symmetries of the fundamental laws of physics. One of the most revolutionary and astonishing discoveries in the history of physics is that objects do not fall down, but toward the center of the earth. Newton's law of gravitation does not pick out a special direction: it is left invariant by rotations.

The history of theoretical physics has witnessed the discoveries of one unexpected symmetry after another. Physics in the late twentieth century consists of the astonishing discovery that as we study Nature at ever deeper levels, Nature displays more and more symmetries.[2]

Consider a set of transformations $T_1, T_2, \cdots$ that leave the laws of physics invariant. Let us first perform the transformation T_j, and then perform the transformation T_i. The transformation that results from this sequence of two transformations is denoted by the "product" $T_i \cdot T_j$. Evidently, if T_i and T_j leave the laws of physics invariant, then the transformation $T_i \cdot T_j$ also leaves the laws of physics invariant.[3]

Here we label the transformations by a discrete index i. In general, the index could also be continuous. Indeed, the transformation could depend on a number of continuous parameters. The classic example is a rotation $R(\theta, \varphi, \zeta)$, which can be completely characterized by three angles, as indicated. For example, in one standard parametrization,[4] the two angles θ and φ specify the unit vector describing the rotation axis, while the angle ζ specifies the angle through which we are supposed to rotate around that axis.

Groups

This discussion rather naturally invites us to abstract the concept of a group.

A group G consists of a set of entities $\{g_\alpha\}$ called group elements (or elements for short), which we could compose together (or more colloquially, multiply together). Given any two

elements g_α and g_β, the product $g_\alpha \cdot g_\beta$ is equal to another element,* say, g_γ, in G. In other words, $g_\alpha \cdot g_\beta = g_\gamma$. Composition or multiplication[5] is indicated by a dot (which I usually omit if there is no danger of confusion). The set of all relations of the form $g_\alpha \cdot g_\beta = g_\gamma$ is called the multiplication table of the group.

Composition or multiplication (we will use the two words interchangeably) satisfies the following axioms:†

1. Associativity: Composition is associative: $(g_\alpha \cdot g_\beta) \cdot g_\gamma = g_\alpha \cdot (g_\beta \cdot g_\gamma)$.
2. Existence of the identity: There exists a group element, known as the identity and denoted by I, such that $I \cdot g_\alpha = g_\alpha$ and $g_\alpha \cdot I = g_\alpha$.
3. Existence of the inverse: For every group element g_α, there exists a unique group element, known as the inverse of g_α and denoted by g_α^{-1}, such that $g_\alpha^{-1} \cdot g_\alpha = I$ and $g_\alpha \cdot g_\alpha^{-1} = I$.

A number of comments follow.

1. Composition is not required to commute.[6] In general, $g_\alpha \cdot g_\beta$ is not equal to $g_\beta \cdot g_\alpha$. In this respect, the multiplication of group elements is, in general, like the multiplication of matrices but unlike that of ordinary numbers.

   A group for which the composition rule is commutative is said to be abelian,‡ and a group for which this is not true is said to be nonabelian.§
2. The right inverse and the left inverse are by definition the same. We can imagine mathematical structures for which this is not true, but then these structures are not groups. Recall (or read in the review of linear algebra) that this property holds for square matrices: provided that the inverse M^{-1} of a matrix M exists, we have $M^{-1}M = MM^{-1} = I$ with I the identity matrix.
3. It is often convenient to denote I by g_0.
4. The label α that distinguishes the group element g_α may be discrete or continuous.
5. The set of elements may be finite (that is, $\{g_0, g_1, g_2, \cdots, g_{n-1}\}$), in which case G is known as a finite group with n elements. (Our friend the jargon guy[7] informs us that n is known as the order of the group.)

Mathematicians[8] of course can study groups on the abstract level without tying g_i to any physical transformation, but in some sense the axioms become clearer if we think of transformations in the back of our mind. For example, $gI = Ig = g$ says that the net effect of first doing nothing and then doing something is the same as first doing something and then doing nothing, and the same as doing something. Existence of the inverse says that the transformations of interest to physics can always[9] be undone.[10]

* This property, known as closure, is sometimes stated as an axiom in addition to the three axioms given below.

† See also appendices 1 and 2.

‡ Named after the mathematician Niels Henrik Abel, one of the founders of group theory.

§ As the reader might have heard, in contemporary physics, the theory of the strong, weak, and electromagnetic interactions are based on nonabelian gauge symmetries. See chapter IX.1.

Examples of groups

To gain a better understanding of what a group is, it is best to go through a bunch of examples. For each of the following examples, you should verify that the group axioms are satisfied.

1. Rotations in 3-dimensional Euclidean space, as already mentioned, form the poster child of group theory and are almost indispensable in physics. Think of rotating a rigid object, such as a bust of Newton. After two rotations in succession, the bust, being rigid, has not been deformed in any way: it merely has a different orientation. Thus, the composition of two rotations is another rotation.

   Rotations famously do not commute. See figure 2.

   Descartes taught us that 3-dimensional Euclidean space could be thought of as a linear vector space, coordinatized with the help of three unit basis vectors $\vec{e}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\vec{e}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and $\vec{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, aligned along three orthogonal directions traditionally named x, y, and z. A rotation takes each basis vector into a linear combination of these three basis vectors, and is thus described by a 3-by-3 matrix. This group of rotations is called $SO(3)$. We shall discuss rotations in great detail in chapter I.3; suffice it to mention here that the determinant of a rotation matrix is equal to 1.

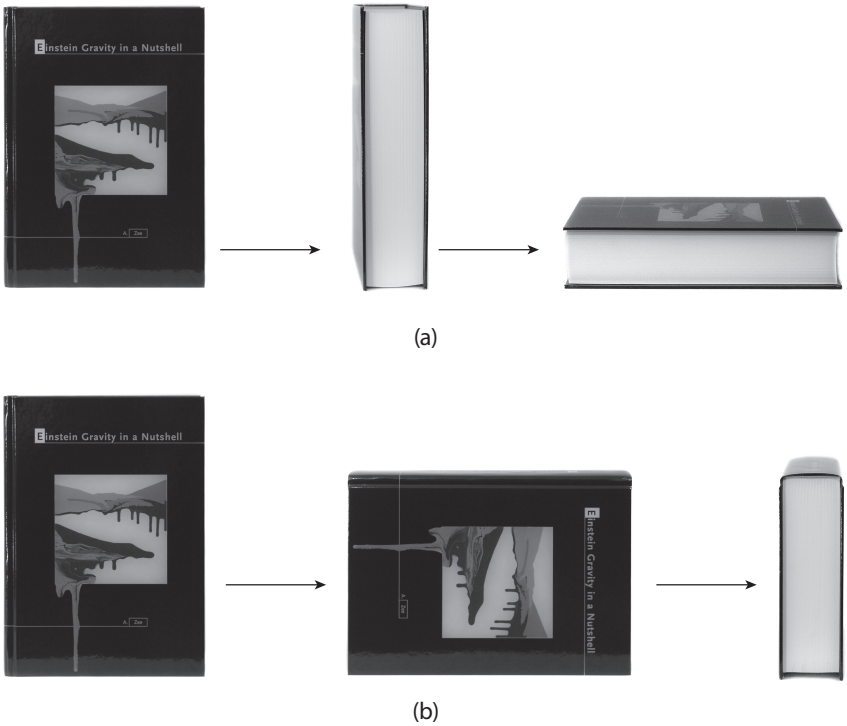

Figure 2

Figure 3

2. Rotations in 2-dimensional Euclidean space, namely a plane, form a group called $SO(2)$, consisting of the set of rotations around an axis perpendicular to the plane. Denote a rotation through angle ϕ by $R(\phi)$. Then $R(\phi_1)R(\phi_2) = R(\phi_1 + \phi_2) = R(\phi_2)R(\phi_1)$. These rotations commute. (See figure 3. I was surprised to discover that this bust of Dirac at St. John's College, Cambridge University, was not nailed down but could be rotated around the vertical axis. The photo depicts my attempt to give the bust a half-integral amount of angular momentum.)

3. The permutation group S_4 rearranges an ordered set of four objects, which we can name arbitrarily, for example, (A, B, C, D) or $(1, 2, 3, 4)$. An example would be a permutation that takes $1 \to 3$, $2 \to 4$, $3 \to 2$, and $4 \to 1$. As is well known, there are $4! = 24$ such permutations (since we have four choices for which number to take 1 into, three choices for which number to take 2 into, and two choices for which number to take 3 into). The permutation group S_n evidently has $n!$ elements. We discuss S_n in detail in chapter I.2.

4. Even permutations of four objects form the group A_4. As is also well known, a given permutation can be characterized as either even or odd (we discuss this in more detail in chapter I.2). Half of the 24 permutations in S_4 are even, and half are odd. Thus, A_4 has 12 elements. The jargon guy tells us that A stands for "alternating."

5. The two square roots of 1, $\{1, -1\}$, form the group Z_2 under ordinary multiplication.

6. Similarly, the three cube roots of 1 form the group $Z_3 = \{1, \omega, \omega^2\}$ with $\omega \equiv e^{2\pi i/3}$.

   Chugging right along, we note that the four fourth roots of 1 form the group $Z_4 = \{1, i, -1, -i\}$, where famously (or infamously) $i = e^{i\pi/2}$.

   More generally, the N Nth roots of 1 form the group $Z_N = \{e^{i2\pi j/N} : j = 0, \cdots, N-1\}$. The composition of group elements is defined by $e^{i2\pi j/N} e^{i2\pi k/N} = e^{i2\pi(j+k)/N}$.

   Quick question: Does the set $\{1, i, -1\}$ form a group?

7. Complex numbers of magnitude 1, namely $e^{i\phi}$, form a group called $U(1)$, with $e^{i\phi_1}e^{i\phi_2} = e^{i(\phi_1+\phi_2)}$. Since $e^{i(\phi+2\pi)} = e^{i\phi}$, we can restrict ϕ to range from 0 to 2π. At the level of physicist rigor, we can think of $U(1)$ as the "continuum limit" of Z_N with $e^{i2\pi j/N} \to e^{i\phi}$ in the limit $N \to \infty$ and $j \to \infty$ with the ratio held fixed $2\pi j/N = \phi$.

8. The addition of integers mod N generates a group. For example, under addition mod 5 the set {0, 1, 2, 3, 4} forms a group: $2 + 1 = 3$, $3 + 2 = 0$, $4 + 3 = 2$, and so on. The composition of the group elements is defined by $j \cdot k = j + k$ mod 5. The identity element I is denoted by 0. The inverse of 2, for example, is 3, of 4 is 1, and so on. The group is clearly abelian.

   Question: Have you seen this group before?

9. The addition of real numbers form a group, perhaps surprisingly. The group elements are denoted by a real number u and $u \cdot v \equiv u + v$, where the symbol $+$ is what an elementary school student would call "add." You can easily check that the axioms are satisfied. The identity element is denoted by 0, and the inverse of the element u is the element $-u$.

10. The additive group of integers is obtained from the additive group of real numbers by restricting u and v in the preceding example to be integers of either sign, including 0.

11. As many readers know, in Einstein's theory of special relativity,[11] the spacetime coordinates used by two observers in relative motion with velocity v along the x-direction (say) are related by the Lorentz transformation (with c the speed of light):

$$\begin{aligned} ct' &= \cosh\varphi\, ct + \sinh\varphi\, x \\ x' &= \sinh\varphi\, ct + \cosh\varphi\, x \\ y' &= y \\ z' &= z \end{aligned} \tag{1}$$

    where the "boost angle" φ is determined by $\tanh\varphi = v$. (In other words, $\cosh\varphi = 1/\sqrt{1-\frac{v^2}{c^2}}$, and $\sinh\varphi = \frac{v}{c}/\sqrt{1-\frac{v^2}{c^2}}$.) Suppressing the y- and z-coordinates, we can describe the Lorentz transformation by

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\varphi & \sinh\varphi \\ \sinh\varphi & \cosh\varphi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{2}$$

    Physically, suppose a third observer is moving at a velocity defined by the boost angle φ_2 relative to the observer moving at a velocity defined by the boost angle φ_1 relative to the first observer. Then we expect the third observer to be moving at some velocity determined by φ_1 and φ_2 relative to the first observer. (All motion is restricted to be along the x-direction for simplicity.) This physical statement is expressed by the mathematical statement that the Lorentz transformations form a group:

$$\begin{pmatrix} \cosh\varphi_2 & \sinh\varphi_2 \\ \sinh\varphi_2 & \cosh\varphi_2 \end{pmatrix} \begin{pmatrix} \cosh\varphi_1 & \sinh\varphi_1 \\ \sinh\varphi_1 & \cosh\varphi_1 \end{pmatrix} = \begin{pmatrix} \cosh(\varphi_1+\varphi_2) & \sinh(\varphi_1+\varphi_2) \\ \sinh(\varphi_1+\varphi_2) & \cosh(\varphi_1+\varphi_2) \end{pmatrix} \tag{3}$$

    The boost angles add.*

12. Consider the set of n-by-n matrices M with determinants equal to 1. They form a group under ordinary matrix multiplication, since as was shown in the review of linear algebra,

* To show this, use the identities for the hyperbolic functions.

the determinant of the product two matrices is equal to the product of the determinants of the two matrices: $\det(M_1M_2) = \det(M_1)\det(M_2)$. Thus, $\det(M_1M_2) = 1$ if $\det(M_1) = 1$ and $\det(M_2) = 1$: closure is satisfied. Since $\det M = 1 \neq 0$, the inverse M^{-1} exists. The group is known as $SL(n, R)$, the special linear group with real entries. If the entries are allowed to be complex, the group is called $SL(n, C)$. (Matrices with unit determinant are called special.)

From these examples, we see that groups can be classified according to whether they are finite or infinite, discrete or continuous. Note that a discrete group can well be infinite, such as the additive group of integers.

Concept of subgroup

In group theory, many concepts are so natural that they practically suggest themselves,[12] for example, the notion of a subgroup. Given a set of entities $\{g_\alpha\}$ that form a group G, if a subset $\{h_\beta\}$ also form a group, call it H, then H is known as a subgroup of G and we write $H \subset G$.

Here are some examples.

1. $SO(2) \subset SO(3)$. This shows that, in the notation $\{g_\alpha\}$ and $\{h_\beta\}$ we just used, the index sets denoted by α and β can in general be quite different; here α consists of three angles and β of one angle.
2. $S_m \subset S_n$ for $m < n$. Permuting three objects is just like permuting five objects but keeping two of the five objects untouched. Thus, $S_3 \subset S_5$.
3. $A_n \subset S_n$.
4. $Z_2 \subset Z_4$, but $Z_2 \not\subset Z_5$.
5. $SO(3) \subset SL(3, R)$.

Verify these statements.

Cyclic subgroups

For a finite group G, pick some element g and keep multiplying it by itself. In other words, consider the sequence $\{g, g^2 = gg, g^3 = g^2g, \cdots\}$. As long as the resulting product is not equal to the identity, we can keep going. Since G is finite, the sequence must end at some point with $g^k = I$. The set of elements $\{I, g, g^2, \cdots, g^{k-1}\}$ forms a subgroup Z_k. Thus, any finite group has a bunch of cyclic subgroups. If k is equal to the number of elements in G, then the group G is in fact Z_k.

Lagrange's theorem

Lagrange[13] proved the following theorem. Let a group G with n elements have a subgroup H with m elements. Then m is a factor of n. In other words, n/m is an integer.

The proof is as follows. List the elements of H: $\{h_1, h_2, \cdots, h_m\}$. (Note: Since H forms a group, this list must contain I. We do not list any element more than once; thus, $h_a \neq h_b$ for $a \neq b$.) Let $g_1 \in G$ but $\notin H$ (in other words, g_1 is an element of G outside H). Consider the list $\{h_1 g_1, h_2 g_1, \cdots, h_m g_1\}$, which we denote by $\{h_1, \cdots, h_m\} g_1$ to save writing. Note that this set of elements does not form a group. (Can you explain why not?)

I claim that the elements on the list $\{h_1 g_1, h_2 g_1, \cdots, h_m g_1\}$ are all different from one another. Proof by contradiction: For $a \neq b$, $h_a g_1 = h_b g_1 \Longrightarrow h_a = h_b$ upon multiplication from the right by $(g_1)^{-1}$ (which exists, since G is a group).

I also claim that none of the elements on this list are on the list $\{h_1, \cdots, h_m\}$. Proof: For some a and b, $h_a g_1 = h_b \Longrightarrow g_1 = h_a^{-1} h_b$, which contradicts the assumption that g_1 is not in H. Note that H being a group is crucial here.

Next, pick an element g_2 of G not in the two previous lists, and form $\{h_1 g_2, h_2 g_2, \cdots, h_m g_2\} = \{h_1, h_2, \cdots, h_m\} g_2$.

I claim that these m elements are all distinct. Again, this proof follows by contradiction, which you can supply. Answer: For $a \neq b$, $h_a g_2 = h_b g_2 \Longrightarrow h_a = h_b$. I also claim that none of these elements are on the two previous lists. Yes, the proof proceeds again easily by contradiction. For example, $h_a g_2 = h_b g_1 \Longrightarrow g_2 = h_a^{-1} h_b g_1 = h_c g_1$, since H is a group, but this would mean that g_2 is on the list $\{h_1, h_2, \cdots, h_m\} g_1$, which is a contradiction.

We repeat this process. After each step, we ask whether there is any element of G left that is not on the lists already constructed. If yes, then we repeat the process and construct yet another list containing m distinct elements. Eventually, there is no group element left (since G is a finite group). We have constructed k lists, including the original list $\{h_1, h_2, \cdots, h_m\}$, namely, $\{h_1, h_2, \cdots, h_m\} g_j$ for $j = 0, 1, 2, \cdots, k-1$ (writing I as g_0).

Therefore $n = mk$, that is, m is a factor of n. QED.

As a simple example of Lagrange's theorem, we can immediately state that Z_3 is a subgroup of Z_{12} but not of Z_{14}. It also follows trivially that if p is prime, then Z_p does not have a nontrivial subgroup. From this you can already sense the intimate relation between group theory and number theory.

Direct product of groups

Given two groups F and G (which can be continuous or discrete), whose elements we denote by f and g, respectively, we can define another group $H \equiv F \otimes G$, known as the direct product of F and G, consisting of the elements (f, g). If you like, you can think of the symbol (f, g) as some letter in a strange alphabet. The product of two elements (f, g) and (f', g') of H is given by $(f, g)(f', g') = (ff', gg')$. The identity element of H is evidently given by (I, I), since $(I, I)(f, g) = (If, Ig) = (f, g)$ and $(f, g)(I, I) = (fI, gI) = (f, g)$. (If we were insufferable pedants, we would write $I_H = (I_F, I_G)$, since the identity elements I_H, I_F, I_G of the three groups H, F, G are conceptually quite distinct.)

What is the inverse of (f, g)? If F and G have m and n elements, respectively, how many elements does $F \otimes G$ have?

Evidently, the inverse of (f, g) is (f^{-1}, g^{-1}), and $F \otimes G$ has mn elements.

Klein's Vierergruppe V

A simple example is given by $Z_2 \otimes Z_2$, consisting of the four elements: $I = (1, 1)$, $A = (-1, 1)$, $B = (1, -1)$, and $C = (-1, -1)$. For example, we have $AB = (-1, -1) = C$. Note that this group is to be distinguished from the group Z_4 consisting of the four elements $1, i, -1, -i$. The square of any element in $Z_2 \otimes Z_2$ is equal to the identity, but this is not true of Z_4. In particular, $i^2 = -1 \neq 1$.

Incidentally, $Z_2 \otimes Z_2$, also known as Klein's Vierergruppe ("4-group" in German) and denoted by V, played an important historical role in Klein's program.

Note that the elements of F, regarded as a subgroup of $F \otimes G$, are written as (f, I). Similarly, the elements of G are written as (I, g). Clearly, (f, I) and (I, g) commute.

The direct product would seem to be a rather "cheap" way of constructing larger groups out of smaller ones, but Nature appears to make use of this possibility. The theory of the strong, weak, and electromagnetic interaction is based on the group* $SU(3) \otimes SU(2) \otimes U(1)$.

A teeny[14] bit of history: "A pleasant human flavor"

Historians of mathematics have debated about who deserves the coveted title of "the founder of group theory." Worthy contenders include Cauchy, Lagrange, Abel, Ruffini, and Galois. Lagrange was certainly responsible for some of the early concepts, but the sentimental favorite has got to be Évariste Galois, what with the ultra romantic story of him feverishly writing down his mathematical ideas the night before a fatal duel at the tender age of 20. Whether the duel was provoked because of the honor of a young woman named du Motel or because of Galois's political beliefs† (for which he had been jailed) is apparently still not settled. In any case, he was the first to use the word "group." Nice choice.

To quote the mathematician G. A. Miller, it is silly to argue about who founded group theory anyway:

> We are inclined to attribute the honor of starting a given big theory to an individual just as we are prone to ascribe fundamental theorems to particular men, who frequently have added only a small element to the development of the theorem. Hence the statement that a given individual founded a big theory should not generally be taken very seriously. It adds, however, a pleasant human flavor and awakens in us a noble sense of admiration and appreciation. It is also of value in giving a historical setting and brings into play a sense of the dynamic forces which have contributed to its development instead of presenting to us a cold static scene. Observations become more inspiring when they are permeated with a sense of development.[15]

* The notation $SU(n)$ and $U(n)$ will be explained in detail later in chapter IV.4.
† Galois was a fervent Republican (in the sense of being against the monarchy, not the Democrats).

While symmetry considerations have always been relevant for physics, group theory did not become indispensable for physics until the advent of quantum mechanics, for reasons to be explained in chapter III.1. Eugene Wigner,[16] who received the Nobel Prize in 1963 largely for his use of group theory in physics, recalled the tremendous opposition to group theory among the older generation (including Einstein, who was 50 at the time) when he first started using it around 1929 or so.[17] Schrödinger told him that while group theory provided a nice derivation of some results in atomic spectroscopy, "surely no one will still be doing it this way in five years." Well, a far better theoretical physicist than a prophet!

But Wigner's childhood friend John von Neumann,[18] who helped him with group theory, reassured him, saying "Oh, these are old fogeys. In five years, every student will learn group theory as a matter of course."[19]

Pauli[20] coined the term "die Gruppenpest" ("that pesty group business"), which probably captured the mood at the time. Remember that quantum mechanics was still freshly weird, and all this math might be too much for older people to absorb.

Multiplication table: The "once and only once rule"

A finite group with n elements can be characterized by its multiplication table,[21] as shown here. We construct a square n-by-n table, writing the product $g_i g_j$ in the square in the ith row and the jth column:

	$\cdots$	g_j	$\cdots$
$\vdots$	$\ddots$		
g_i		$g_i g_j$	
$\vdots$			$\ddots$

A simple observation is that, because of the group properties, in each row any group element can appear once and only once. To see this, suppose that in the ith row, the same group element appears twice, that is, $g_i g_j = g_i g_k$ for $j \neq k$. Then multiplying by g_i^{-1} from the left, we obtain $g_j = g_k$, contrary to what was assumed. It follows that each of the n elements must appear once to fill up the n slots in that row. We might refer to this as the "once and only once rule."

The same argument could be repeated with the word "row" replaced by "column," of course.

For n small, all possible multiplication tables and hence all possible finite groups with n elements can readily be constructed. Let us illustrate this for $n = 4$. For pedagogical reasons, we will do this in two different ways, one laborious,* the other "slick."

* An undergraduate in my class advised me to include also this laborious way as being the more instructive of the two ways. I agree with him that textbooks tend to contain too many slick proofs.

Finite groups with four elements: The slow way

First, we proceed very slowly, by brute force. Call the four elements I, A, B, and C.

1. By definition of the identity, the first row and first column can be filled in automatically:

	I	A	B	C
I	I	A	B	C
A	A			
B	B			
C	C			

2. We are to fill in the second row with I, B, and C. The first entry in that row is A^2. There are two possible choices: choice (a): $A^2 = B$, or choice (b): $A^2 = I$. (You might think that there is a third choice, $A^2 = C$, but that is the same as choice (a) upon renaming the elements. What you call C I will call B.)

   Let us now follow choice (a) and come back to choice (b) later.

3. The multiplication table now reads

	I	A	B	C
I	I	A	B	C
A	A	B	2	3
B	B	4	5	6
C	C			

   where for your and my convenience I have numbered some of the boxes yet to be filled in.

4. We have to put C and I into boxes 2 and 3. But we cannot put C into box 3, since otherwise the fourth column will break the "once and only once rule": C would appear twice:

	I	A	B	C
I	I	A	B	C
A	A	B	C	I
B	B	4	5	6
C	C			

5. Again by the "once and only once rule," box 4 can only be C or I. The latter choice would mean $BA = I$ and hence $B = A^{-1}$, but we already know from the second row of the multiplication table that $AB = C \neq I$. Thus, box 4 can only be C. Hence box 5 is I, and 6 is A.

6. Finally, the last three blank entries in the fourth column are fixed uniquely by the "once and only once rule." We obtain[22]

	I	A	B	C
I	I	A	B	C
A	A	B	C	I
B	B	C	I	A
C	C	I	A	B

Now that we have the multiplication table, we know everything about the group, and we can ask: What group is this? From the second row, we read off $A^2 = B$, $A^3 = AA^2 = AB = C$, $A^4 = AA^3 = AC = I$. The group is Z_4. Interestingly, we don't even have to finish constructing the entire multiplication table. In this simple case, by the time we had filled in the second row, we could have quit.

The rest of the table, however, provides us with a lot of consistency checks to ensure that we have not messed up. For example, from the last row, we have $CB = A$. But we know from the second row that $B = A^2$ and $C = A^3$, and hence the statement $CB = A$ says that $A^3A^2 = A^5 = A$, showing that indeed $A^4 = I$.

We now go back to choice (b): $A^2 = I$, so that

	I	A	B	C
I	I	A	B	C
A	A	I	2	3
B	B	4	5	6
C	C	7	8	9

1. We are to fill boxes 2 and 3 with C and B. By the "once and only once rule" in the third and fourth columns, these boxes can only be C and B in that order.

2. By the same reasoning, we can only fill boxes 4 and 7 with C and B, respectively. We thus obtain

	I	A	B	C
I	I	A	B	C
A	A	I	C	B
B	B	C		
C	C	B		

3. Now it looks like we could fill in the four remaining empty boxes with either

I	A
A	I

or

A	I
I	A

But the two choices amount to the same thing. We simply rename B and C. Thus, we obtain

	I	A	B	C
I	I	A	B	C
A	A	I	C	B
B	B	C	I	A
C	C	B	A	I

Again, what group is this? It is just $Z_2 \otimes Z_2$: $A^2 = I$, $B^2 = I$, $C = AB = BA$ (and hence also $C^2 = I$).

A quick way: Construct the cyclic subgroups

Here is an alternative to this laborious procedure of constructing the multiplication table step by step. We use the earlier observation that in a finite group, if we keep multiplying an element by itself, we will reach the identity I.

Given a group G of four elements $\{I, A, B, C\}$, we keep multiplying A by itself. If $A^4 = I$, then $G = Z_4$. By Lagrange's theorem, the possibility $A^3 = I$ is not allowed. If $A^2 = I$, then we multiply B by itself. Either B^2 or B^4 equals I. The latter is ruled out, so the only possibility is that $B^2 = I$, and $AB = BA = C$. Then $G = Z_2 \otimes Z_2$, with the four elements represented by $(1, 1)$, $(1, -1)$, $(-1, 1)$, and $(-1, -1)$.

If you are energetic and driven, you could try to construct all possible finite groups with n elements, and see how large an n you could get to.[23] A quick hint: It's easy if n is prime.

Presentations

For large groups, writing down the multiplication table is clearly a losing proposition. Instead, finite groups are defined by their properties, as in the examples listed above, or by

their presentations,[24] which list the elements (sometimes called generators) from which all other elements can be obtained by group multiplication, and the essential relations the generators satisfy. Thus, in a self-evident notation, the groups Z_4 and $Z_2 \otimes Z_2$ are defined by their presentations as follows:

$$Z_4 : \langle A | A^4 = I \rangle \tag{4}$$

$$Z_2 \otimes Z_2 : \langle A, B \, | A^2 = B^2 = I, \, AB = BA \rangle \tag{5}$$

The two groups are clearly distinct. In particular, Z_4 contains only one element that squares to I, namely A^2.

Homomorphism and isomorphism

A map $f : G \to G'$ of a group G into the group G' is called a homomorphism if it preserves the multiplicative structure of G, that is, if $f(g_1) f(g_2) = f(g_1 g_2)$. Clearly, this requirement implies that $f(I) = I$ (more strictly speaking, the identity of G is mapped to the identity of G'). A homomorphism becomes an isomorphism if the map is one-to-one and onto.

Now we can answer the question posed earlier: the additive group of integers mod N is in fact isomorphic* to Z_N.

For a more interesting example, consider $Z_2 \otimes Z_4$. We use the additive notation here and thus write the elements as (n, m) and compose them according to $(n, m) \cdot (n', m') = (n + n' \bmod 2, m + m' \bmod 4)$. We start with $(0, 0)$ and add $(1, 1)$ repeatedly: $(0, 0) \xrightarrow{+(1,1)} (1, 1) \to (0, 2) \to (1, 3) \to (0, 4) = (0, 0)$; we get back to where we started. Next, we start with $(0, 1)$ and again add $(1, 1)$ repeatedly: $(0, 1) \xrightarrow{+(1,1)} (1, 2) \to (0, 3) \to (1, 0) \to (0, 1)$, getting back to where we started. Thus we can depict $Z_2 \otimes Z_4$ by a rectangular 2-by-4 discrete lattice on a torus (see figure 4).

Now we come in for a bit of a surprise. Consider $Z_2 \otimes Z_3$ consisting of (n, m), which we compose by $(n + n' \bmod 2, m + m' \bmod 3)$. Again, we start with $(0, 0)$ and add $(1, 1)$ repeatedly: $(0, 0) \to (1, 1) \xrightarrow{+(1,1)} (2, 2) = (0, 2) \to (1, 3) = (1, 0) \to (2, 1) = (0, 1) \to (1, 2) \to (2, 3) = (0, 0)$. We are back where we started! In the process, we cycled through all six elements of $Z_2 \otimes Z_3$. We conclude that the six elements $(0, 0)$, $(1, 1)$, $(0, 2)$, $(1, 0)$, $(0, 1)$, and $(1, 2)$ describe Z_6.

Thus, $Z_2 \otimes Z_3$ and Z_6 are isomorphic; they are literally the same group. Note that this phenomenon, of a possible isomorphism between $Z_p \otimes Z_q$ and Z_{pq}, does not require p and q to be prime, only relatively prime. (Consider the example of $Z_4 \otimes Z_9$.)

As another example of isomorphism, the groups $SO(2)$ and $U(1)$ introduced earlier in the chapter are isomorphic. The map $f : SO(2) \to U(1)$ is defined simply by $f(R(\phi)) = e^{i\phi}$.

* That the additive group of integers mod N is also isomorphic to the multiplicative group Z_n foreshadows the confusion some students have between the addition and multiplication of angular momenta in quantum mechanics. We discuss this later in chapter IV.3.

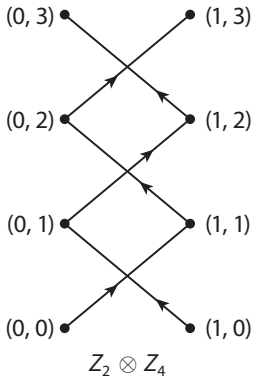

Figure 4

Appendix 1: Weakening the axioms

Two of the three axioms that define a group can in fact be weakened to the following:

2′. Existence of the left identity: A left identity I exists, such that for any element g, $Ig = g$.

3′. Existence of a left inverse: For any element g, there exists an element f, such that $fg = I$.

We now show that these imply axioms 2 and 3 given in the text. In other words, given the left identity and the left inverse, we are guaranteed that the right identity and the right inverse also exist.

Take the left inverse f of g. By 3′, there exists an element k, such that $kf = I$. Multiplying this by g from the right, we obtain $(kf)g = Ig = g = k(fg) = kI$, where the second equality is due to 2′, the third equality to associativity, and the fourth equality to 3′. Therefore $g = kI$. We want to show that $k = g$.

To show this, let us multiply $g = kI$ by I from the right. We obtain $gI = (kI)I = k(II) = kI = g$, where the second equality is due to associativity, and the third equality to 2′, since I also qualifies as "any element." Thus, $gI = g$, so that I is also the right identity. But if I is also the right identity, then the result $g = kI$ becomes $g = k$. Multiplying by f from the right, we obtain $gf = kf = I$. Therefore, the left inverse of g, namely f, is also the right inverse of g.

Appendix 2: Associativity

Mathematically, the concept of a group is abstracted from groups of transformations. To physicists, groups are tantamount to transformation groups. In fact, if we are allowed to think of group elements as acting on a set of things $\mathcal{S} = \{p_1, p_2, \cdots\}$, we can prove associativity. The "things" could be interpreted rather generally. For the geometrical examples given in this chapter, p_i could be the points in, for example, a triangle. Or for applications to fundamental physics, p_i could be some physical law as known to a particular observer, for example, an inertial observer in discussions of special relativity.

Suppose the group element g takes $p_1 \to p_1'$, $p_2 \to p_2'$, $\cdots$, so that the things in $\mathcal{S}$ are rearranged (as, for example, when a triangle is rotated). Suppose the group element g' takes $p_1' \to p_1''$, $p_2' \to p_2''$, $\cdots$, and the group element g'' takes $p_1'' \to p_1'''$, $p_2'' \to p_2'''$, $\cdots$, and so on.

Now consider the action of $g''(g'g)$ on $\mathcal{S}$. The element $g'g$ takes p_j to p_j'', and then the element g'' takes p_j'' to p_j'''. Compare this with the action of $(g''g')g$ on $\mathcal{S}$. The element g takes p_j to p_j', and then the element $g''g'$ takes p_j' to p_j'''. The final result is identical, and associativity is proved.

Most physicists I know would probably regard this kind of fundamental proof as painfully self-evident.

Appendix 3: Modular group

The modular group has become important in several areas of physics, for example, string theory and condensed matter physics. Consider the set of transformations of one complex number into another given by

$$z \to \frac{az+b}{cz+d} \tag{6}$$

with a, b, c, and d integers satisfying $ad - bc = 1$. The transformation (6) can be specified by the matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{with } \det M = 1 \tag{7}$$

Clearly, M and $-M$ correspond to the same transformation in (6).

In the text, I introduced you to $SL(n, R)$, the special linear group of n-by-n matrices with real entries, and $SL(n, C)$, the special linear group of n-by-n matrices with complex entries. The matrices in (7) define the group $SL(2, Z)$, the special linear group of 2-by-2 matrices with integer entries.* The group that results upon identifying M and $-M$ in $SL(2, Z)$ is known as $PSL(2, Z)$ (the letter P stands for "projective"), otherwise known as the modular group.

The transformation in (6) can be generated by repeatedly composing (that is, multiplying together) the two generating transformations

$$S : z \to -\frac{1}{z} \tag{8}$$

and

$$T : z \to z + 1 \tag{9}$$

They correspond to the matrices $S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, respectively.

Using the language of presentation introduced in the text, we can write

$$PSL(2, Z) : \langle S, T \mid S^2 = I,\ (ST)^3 = I \rangle \tag{10}$$

Incidentally, the modular group can be generalized to the triangular group $\mathcal{T}$, denoted by $(2, 3, n)$ and presented by

$$\mathcal{T} : \langle S, T \mid S^2 = I,\ (ST)^3 = I,\ T^n = I \rangle \tag{11}$$

The modular group is thus sometimes written as $(2, 3, \infty)$.

Exercises

1 The center of a group G (denoted by Z) is defined to be the set of elements $\{z_1, z_2, \cdots\}$ that commute with all elements of G, that is, $z_i g = g z_i$ for all g. Show that Z is an abelian subgroup of G.

2 Let $f(g)$ be a function of the elements in a finite group G, and consider the sum $\sum_{g \epsilon G} f(g)$. Prove the identity $\sum_{g \epsilon G} f(g) = \sum_{g \epsilon G} f(gg') = \sum_{g \epsilon G} f(g'g)$ for g' an arbitrary element of G. We will need this identity again and again in chapters II.1 and II.2.

* In mathematics, Z denotes the set of all integers, of either sign, including 0.

3 Show that $Z_2 \otimes Z_4 \neq Z_8$.

4 Find all groups of order 6.

Notes

1. See Fearful.
2. See parts VII and VIII.
3. We go into this in detail in chapter III.3.
4. See chapter IV.7.
5. Of course, we could also be more abstract and say that a group G is a structure endowed with the map $(G, G) \to G$ and so on and so forth.
6. In Strange Beauty, the biography of Murray Gell-Mann by G. Johnson, the following explanation about commutation is mentioned. When Gell-Mann was admitted only to MIT rather than the graduate school of his choice, he resolved to kill himself. But then he realized that killing himself and attending MIT do not commute, and so decided that he should go to MIT first and kill himself later, rather than the other way around.
7. He is one of several characters that populate my previous books Quantum Field Theory in a Nutshell and Einstein Gravity in a Nutshell. Hereafter QFT Nut and G Nut, respectively.
8. In the late nineteenth century, mathematicians felt that, with group theory, they had finally invented something of no use to the physicists. See p. v in R. Gilmore, Lie Groups, Lie Algebras, and Some of Their Applications.
9. Note the conceptual distinction between transformation and invariance. For example, the laws governing the weak interaction are famously not invariant under the interchange of left and right (known as a parity transformation P). But, regardless of whether a given law is invariant under parity, we still have $P \cdot P = I$.
10. This unfortunately is not true of many transformations in everyday life, such as cooking and aging.
11. See, for example, G Nut.
12. I once had a math professor who spoke of self-proving theorems. In the same sense, there are self-suggesting concepts.
13. Lagrange fell into a deep depression in his old age. Fortunately for him, the daughter of Lemonnier, an astronomer friend of Lagrange's, managed to cheer him up. Almost forty years younger than Lagrange, the young woman offered to marry him. Soon Lagrange was productive again. "Mathematicians Are People, Too," by L. Reimer and W. Reimer, p. 88.
14. "Teeny bit of history," because you can easily read your fill on the web.
15. "The Founder of Group Theory" by G. A. Miller, American Mathematical Monthly 17 (Aug–Sep 1910), pp. 162–165. http://www.jstor.org/stable/2973854.
16. My senior colleague Robert Sugar, who took a course on group theory at Princeton from Wigner, told me the following story. On the first day, Wigner asked the students whether they knew how to multiply matrices. Given Wigner's reputation of delivering long dull discourses, the students all said yes of course, and in fact, as graduate students at Princeton, they all knew how to do it. But Wigner was skeptical and asked a student to go up to the blackboard and multiply two 2-by-2 matrices together. The guy did it perfectly, but unfortunately, Wigner used a convention opposite to what was (and still is) taught in the United States. Wigner was convinced that the students did not know how to multiply matrices, and proceeded to spend a week tediously explaining matrix multiplication. If you look at the English edition of Wigner's group theory book, you would read that the translator had, with Wigner's permission, reversed all of his conventions.
17. The stories Wigner told about the early days of group theory used here and elsewhere in this bookare taken from The Recollections of Eugene P. Wigner as told to Andrew Szanton, Plenum Press, 1992.
18. As you might have heard, the four Hungarians, Leo Szilard, Eugene Wigner, John von Neumann, and Edward Teller, all Jewish, formed a legendary group that had major impact on physics. Listed here in order of age, they were born within 10 years of one another. Wigner considered himself to be the slowest of the four, and anecdotal evidence suggests that this assessment is not due to exaggerated modesty; yet he is the only one of the four to have received a Nobel Prize.
19. Well, not quite—not even close.

20. This surprises me, since one of Pauli's famous contributions involves group theory. See the interlude to part VII. From what I have read, Pauli was brilliant but mercurial and moody, and always ready for a good joke.
21. As a child you memorized the standard 9-by-9 multiplication table; now you get the chance to construct your own.
22. Conspiracy nuts might notice that the acronym CIA appears not once, but four times, in this table.
23. Mathematicians have listed all possible finite groups up to impressively large values of n.
24. As in the rather old-fashioned and formal "May I present [Title] So-and-so to you?"