Chapter 3
Symmetry

Weyl's classic book on symmetry conveys the idea that the notion of symmetry is not just an art or an invention of the mind, but part of the observational structure of nature [59]. However, the awareness of the importance of symmetry in physics became clear only after the debate on the negative result of the Michelson–Morley experiment and the subsequent interpretation of the relative motion between observers and observables given by Einstein. This interpretation led us to the emergence of the Poincaré symmetry. From then on, the structure of *Lie symmetry* has become the essential tool to the understanding of the fundamental interactions.

3.1 Groups and Subgroups

A group G is a set composed of elements $a, b, c, \ldots$, endowed with a closed operation (generically denoted by $*$) such that

(a) The operation is associative: $a * (b * c) = (a * b) * c$;
(b) There is a neutral or identity element 1: $a * 1 = 1 * a = a$; and
(c) For each element $a \in G$, there is an inverse element denoted by a^{-1} such that $a^{-1} * a = a * a^{-1} = 1$.

For an Abelian or commutative group we also have $a * b = b * a$. The number of elements in a group is called the *order of the group*. A group of infinite order has infinite elements.

A subset $H \subset G$ is a *subgroup* of G, when its elements form a group *with the same operation of* G. If $H \subset G$, $H \neq G$, then H is a *proper subgroup* of G. It is easy to see that the identity of G must be also contained in all subgroups of G.

Like a manifold, a group can be parameterized by a set of real numbers $(\theta^1, \ldots, \theta^N)$, given by 1:1 maps or charts $X : G \rightarrow I\!R^N$ such that for each element $r \in G$ we have an element of $I\!R^N$

$$X(r) = (\theta^1, \ldots, \theta^N)$$

and conversely, given a point in $I\!R^N$ we obtain an element of G

M.D. Maia, *Geometry of the Fundamental Interactions*,
DOI 10.1007/978-1-4419-8273-5_3,

$$r = X^{-1}(\theta^1, \ldots, \theta^N) = X^{-1}(\theta)$$

The *dimension of a group* is the maximum number N of independent parameters required to describe any element of the group.

Given two elements $r = X_1(\theta)$ and $s = Y^{-1}(\theta')$, the group composition gives $r * s = X^{-1}(\theta) * Y^{-1}(\theta') = t = Z^{-1}(\theta'') \in G$. The parameters θ'' must then be related to θ and θ' as

$$\theta'' = Z \circ (X^{-1}(\theta) * Y^{-1}(\theta')) = f(\theta, \theta') \tag{3.1}$$

Since these charts cover the whole group, they form an *atlas* similar to the case of differentiable manifolds. Then the above condition (3.1) must be satisfied for all elements of the group (such condition does not exist in differentiable manifolds).

When the parameters vary continuously within a given interval on $I\!R^N$ and (3.1) is a homeomorphism we have a *continuous group*. This is less demanding than the differentiable manifold structure where the relation between the parameters (the coordinates) is a diffeomorphism. However, later on we shall be using an even stronger condition imposed by Lie, where (3.1) is required to be an analytic function.

For notational simplicity, from now on we will omit the $*$ operation and write it simply as a product.[1] Thus $r * s$ is written simply as rs.

Definition 3.1 (Cosets and Normal Subgroups) Consider a subgroup of $H \subset G$ and r a specific element of G. Then the set denoted by

$$rH = \{rx \mid x \in H\}$$

is called the *left coset* of H. Similarly we may define the *right coset* of H denoted by Hr. It follows from this definition that $rH \neq H$ because r is not necessarily in H. However, rH necessarily contains r because as a subgroup H contains the identity 1. Hence $r = r1 \in rH$.

Given two left cosets (or right cosets) of a subgroup H in G, aH and bH, if they possess a common element then they are necessarily identical.

Indeed, consider the left cosets $A = aN$ and $b = bN$ and let x be a common element belonging to A and B. Then we may write $x = ar$, $x = bs$, $r, s \in H$, it follows that $ar = bs$ and therefore, $a = bsr^{-1} = bt$, $t = sr^{-1} \in H$. Consequently,

$$aH = \{am \mid m \in H\} = \{btm \mid m, t \in H\} = \{bn \mid n \in H\} = bH$$

A subgroup N in G such that its left and right cosets are equal is called an *invariant* (or *normal*) subgroup of G. From the defining condition $aN = Na$, it follows that

[1] Except for additive groups where the operation is a sum and the neutral element is zero.

$$N = a^{-1}Na = \{q = a^{-1}xa \,|\, x \in N,\ a \in G\}$$

In other words, the elements of an invariant subgroup belong to a class of equivalence where they differ by an equivalence relationship $q \sim x$ defined by $q = a^{-1}xa$.

An interesting property is that if $A = aN$ and $B = bN$ are cosets of an invariant subgroup N, then the set $C = AB = \{xy \in G \,|\, x \in A$ and $y \in B\}$ is also a coset of G. Indeed, writing $x = ap$ and $y = bq,\ p, q \in N$, it follows that the elements of C have the form $xy = apbq$. Since N is an invariant subgroup the left and right cosets of N are identical. Hence, if $p, q \in N$ then $aga^{-1} = r$ and $bab^{-1} = a$ where $r, s \in s$. Therefore, $ap = cp$ and $bs = sb$ and $af = apbq = apsb$. However, $p, s \in N,\ ps \in N$, and using again the fact that N is invariant, $psb = bm,\ m \in N$. Consequently, $xy = abm = cm,\ c = ab$, which implies that xy belongs to a left coset of N, $C = AB = cN$.

The above result suggests the construction of a product operation between cosets of a group G as follows: Given two left cosets A and B defined by the same invariant subgroup N in G, then $C = AB$ is also a left coset cN where $c = ab,\ A = aN$ and $B = bN$.

It can be easily seen that this product defines a group, where the identity element is N:

$$A = AN = \{xy = cm \,|\, c = a1,\ m \in N\} = \{z = am \,|\, m \in\} = A$$

the inverse of $A = aN$ is $A^{-1} \stackrel{\text{def}}{=} a^{-1}N$:

$$C = AA^{-1} = \{xy = cm | c = aa^{-1} = 1,\ m \in N\} = \{z = m | m \in N\} = N.$$

Finally, the product of cosets is associative: If $A = aN,\ B = bN,\ C = cN$, then $(AB)C = (ab)cN = a(bc)N = A(BC)$.

The set of all cosets of an invariant subgroup N, like $A = aN$, defines a group, called the quotient group G/N, with respect to the above defined coset product.

3.2 Groups of Transformations

Groups can be studied by themselves as abstract groups. On the other hand, symmetry groups are *transformation groups*, whose elements are *operators acting on a space or manifold*. We shall be dealing mostly with transformation groups, separated in two cases which are of immediate interest to field theory and to the fundamental interactions. They are *groups of coordinate transformations* acting on the coordinates of a space–time manifold and the *groups of field transformations* acting on field variables.

The groups of *coordinate transformations* act on the coordinate spaces of the manifolds, changing a given coordinate system to another as

$$x'^i = f^i(x^1, x^2, \ldots, x^n, \theta^1, \theta^2, \ldots, \theta^N) = f^i(x^\mu, \theta^a) \qquad (3.2)$$

where in the abbreviated notation x^μ are the old coordinates and θ^a are the parameters of the group.

A simple example of a coordinate transformation group is given by a group of *linear operators* acting on the parameter space $I\!R^3$ of some three-dimensional manifold. If (e_μ) is an arbitrary basis of $I\!R^3$, then the action of the group on that space can be obtained by the action of the group on that basis. For a linear operator $r \in G$ the result is a linear combination of the same basis elements:

$$r(e_i) = r_i^j e_j$$

The quantities r_i^j define a matrix in that basis which represents the group action.

Definition 3.2 (Representations of a Group) A representation of an abstract group G by a transformation group G' is a *homomorphism* $\mathscr{R} : G \to G'$. In the product notation the homomorphism writes as

$$\mathscr{R}(xy) = \mathscr{R}(x)\mathscr{R}(y) \qquad (3.3)$$

Of particular importance is a *linear representation of a group* G, which is the homomorphism

$$\mathscr{R} : G \to G'$$

where G' is a *group of linear operators* acting on some vector space V, called the *representation space*.

Therefore, for each element r of the group G there is a corresponding operator $\mathscr{R}(r)$ acting linearly on the *representation space*. Since $\mathscr{R}$ is a homomorphism,

$$\mathscr{R}(rs) = \mathscr{R}(r)\mathscr{R}(s)$$

From the properties of groups it follows that $\mathscr{R}(r) = \mathscr{R}(r1) = \mathscr{R}(r)\mathscr{R}(1)$. Hence $\mathscr{R}(1) = 1$ and $\mathscr{R}(r^{-1}) = \mathscr{R}(r)^{-1}$.

Therefore, to find a linear representation of a given group, we need in the first place to define a representation space where a transformation group G' acts linearly. Then determine how the group G' acts on that space. Finally, choose a basis of the representation space.

Generally speaking, given one basis $\{\eta_i\}$ in the representation space, the linear representation is defined by the coefficients $\mathscr{R}_j^i$ in the operation

$$\mathscr{R}(r)\eta_i = \mathscr{R}_i^j(r)\,\eta_j$$

Note also that a homomorphism between groups is not necessarily a 1:1 map. In particular we may have several objects in G' corresponding to the identity element of G. The set of all such elements is called the *kernel of the representation*. Denoting this kernel by K, the above definition says that $\mathscr{R}(K) = 1$.

A *faithful linear representation* of a group is a linear representation which is 1:1. That is,

$$\mathscr{R}(r) = \mathscr{R}(s) \iff r = s$$

In this case the kernel contains only the identity element: $K = \{1\}$.

3.3 Lie Groups

A *continuous group* is such that its elements vary continuously with its parameters. The relation between parameters (3.1) can be just a homeomorphism: continuous with an inverse which is also continuous.

Like in a manifold a continuous group may have an induced topology from its parameter space, so that the operations of limits and derivatives can be defined. From this topology it is easy to infer that we may define continuous curves on a continuous group G as a continuous map $\alpha : I\!R \to G$, with tangent vector at a point $r = \alpha(t_0) \in G$, given by $\alpha'(t) = d\alpha/dt\rfloor_{t_0}$ as long as the relation between the parameters (3.1) remains valid.

Consequently we may define on a continuous group some topological properties and classify them according to topological characteristics such as

1. When any two elements of a continuous group G can be connected by any continuous curve or by a continuous sequence of segments of continuous curves, then G is called a *connected group*.
2. A group G is *multiple connected* when there are multiple curves connecting any two elements of G, but they cannot be continuously deformed into one another.
3. A group is compact when each of its parameters θ^a varies in a closed and limited interval.

Definition 3.3 (Lie Group) A continuous group G is a *Lie group* when the composition between the parameters (3.1) is analytic in the sense that $f(\theta, \theta')$ can be represented by converging positive power series [60]. We will see the relevance of this condition when discussing Lie's theorem.

A coordinate transformation produced by a Lie group acting on a differentiable manifold can be written as

$$x'^{\mu} = f^{\mu}(x^{\nu}, \theta^a) \tag{3.4}$$

where $f^{\mu} = x'^{\mu}$ are differentiable functions of the coordinates x^{μ}, but as a consequence of the analyticity of (3.1), they are analytic functions of the parameters θ^a. The local inverse transformation can be either postulated or derived from the condition that the transformation (3.4) is also *regular* in x, that is, when the Jacobian matrix

$$J(f) = \left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right)$$

is non-singular.

A natural linear representations of the Lie group of a coordinate transformations on a manifold is given by the action of the group in the tangent and cotangent spaces, using, respectively, the coordinate basis $\{e_\mu = \partial/\partial x^\mu\}$ and its dual $\{e^\mu = dx^\mu\}$. Taking an arbitrary element $r \in G$, its action on the coordinates also changes the tangent basis as

$$e'_\mu = \frac{\partial}{\partial x'^{\mu}} = \mathscr{R}(r)e_\mu = \sum \mathscr{R}(r)_\mu{}^\nu e_\nu = \sum \mathscr{R}(r)_\mu{}^\nu \frac{\partial}{\partial x^\nu}$$

where $\mathscr{R}(r)_\mu{}^\nu$ denote the matrix elements of the linear representation defined in the tangent space.

Similarly, we obtain the dual representation of the same group using the cotangent space, with the dual coordinate basis $\{e^\mu = dx^\mu\}$. In this basis the linear representation is given by

$$e'^{\mu} = dx'^{\mu} = \sum \mathscr{R}^*(r)^\mu{}_\nu e^\nu = \sum \mathscr{R}^*(r)^\mu{}_\nu dx^\nu$$

where $\mathscr{R}^*(r)_\mu{}^\nu$ denote the matrix elements of the dual linear representation defined in the cotangent space.

More generally we may consider the group G acting on the fibers $\mathscr{V}_p$ defined on an arbitrary vector bundle $(\mathscr{M}, \pi, \mathscr{V})$, where each fiber $\mathscr{V}_p$ is a vector space in which a generic field Ψ, is defined.

Like in the tangent spaces, the action of the group on these fields can be defined by its action on a field basis of these spaces. For example, denoting a basis in one such space by $\{\eta_i\}$, then the group action on a vector $\mathscr{V}_p$ can be determined by its action on that basis as

$$\eta'_i = \mathscr{R}(r)e_i = \sum \mathscr{R}(r)_i{}^j \eta_j \tag{3.5}$$

and a similar linear representation can be defined in the dual basis of the dual fiber $\mathscr{V}^*_p$.

The following table shows some examples of Lie symmetry groups which are relevant for the current development of the theory of the fundamental interactions. Some of these groups will be also discussed in the next sections. For more on exceptional groups see, e.g., [61].

Group	Name	Group elements	Parameters
	Galilean group	3 rotations + 3 boosts + 3 translations + 1 time scale	10
	General Galilean group	3 Rotations + 3 general boosts + 1 time scale + Newton's potential gauge	10
P_4	Poincaré group	6 Pseudo-rotations + 4 translations	10
C_0	Conformal group	Poincaré subgroup + SCT[a] + dilatations + inv.	15
dS_n	deSitter group	Pseudo-rotations on n-dimensional positive sphere	$n(n+1)/2$
AdS_n	Anti-deSitter groups	Pseudo-rotations on n-dimensional negative sphere	$n(n+1)/2$
$GL(N, \mathbb{R})$	Real linear group	Real $N \times N$ matrices	N^2
$SL(N)$	Special linear group	Complex $N \times N$ matrices with determinant 1	$2(N^2-1)$
$SL(N)$	Unimodular group	Real $N \times N$ matrices	N^2-1
$U(N)$	Unitary group	Unitary matrices	N^2
$SU(N)$	Special unitary group	Unitary matrices with determinant 1	N^2-1
$SO(N)$	Special orthogonal group	Real orthogonal matrices with determinant 1	$N(N-1)/2$
G_2	Smallest exceptional group	Automorphisms of octonions	14
E_8	Largest exceptional group	The symmetry group of its Lie algebra	248

[a]Special conformal transformations

3.4 Lie Algebras

The relevance of continuous groups for the study of symmetries is that they allow us to consider infinitesimal transformations defined by when the parameters are small in the presence of unity. As before, we start with the simpler case of a group of coordinate transformations on a manifold $\mathscr{M}$.

3.4.1 *Infinitesimal Coordinate Transformations*

Consider a coordinate transformation described in (3.4) $x'^{\mu} = f^{\mu}(x^{\nu}, \theta^{a})$, followed by a second transformation to another set of coordinates close to x'^{μ}. By the continuity of the group, the parameters of this second transformation must correspond to a small deviation from θ. That is,

$$x''^{\mu} = f^{\mu}(x^{\nu}, \theta^{a} + \delta\theta^{a})$$

Next, expand this function in a Taylor series around $\delta\theta^{\mu} = 0$. Keeping only the first power of $\delta\theta$, we obtain

$$x''^{\mu} = f^{\mu}(x^{\mu}, \theta^{a}) + \sum \frac{\partial f^{\mu}(x, \theta^{a} + \delta\theta^{a})}{\partial \theta^{a}}\Bigg\rfloor_{\delta\theta=0} \delta\theta^{a} = x'^{\mu} + a^{\mu}_{a}(x)\delta\theta^{a} = x'^{\mu} + \xi^{\mu}$$

where we have denoted $\xi^{\mu} = a^{\mu}_{a}(x')\delta\theta^{a}(x, \theta)$, called the *infinitesimal descriptor* of the transformation. As a consequence of (3.1), these functions are also analytic in θ^{a}.

The array a^{μ}_{a} depend on x and θ so that the inverse transformation exists only if it has rank equal to the smallest value between N and n. Simplifying, we may drop the excess primes to write the above infinitesimal coordinate transformation as

$$x'^{\mu} = x^{\mu} + \xi^{\mu} \tag{3.6}$$

To obtain the transformations of fields consider first the infinitesimal coordinate transformation on a differentiable real function F on the manifold

$$dF = \frac{\partial F}{\partial x^{\mu}} dx^{\mu} = \frac{\partial F}{\partial x^{\mu}} a^{\mu}_{a}(x)\delta\theta^{a} = \delta\theta^{a} X_{a} F$$

where we have introduced a linear operator acting on the space of all such differentiable functions on $\mathscr{M}$ by

$$X_{a} = \sum a^{\mu}_{a}(x) \frac{\partial}{\partial x^{\mu}} \tag{3.7}$$

Using these operators the infinitesimal variation of the function can be expressed as

$$F' = F + dF = (1 + \sum \delta\theta^{a} X_{a})F$$

In particular, taking F to be any coordinate x^{μ} we obtain our previous infinitesimal coordinate transformation (3.6).

The linear operators (3.7) generate an N-dimensional vector space with the operations of sum and multiplication by numbers given by

$$(aX_{a} + bY_{a})f = aX_{a}f + bY_{a}f, \; a, b \in I\!R$$

Indeed, suppose that there are constants $c_{a} \in I\!R$, such that $\sum c_{a}X_{a} = 0$. Applying this to x^{μ}, we obtain

$$\sum c_{a}X_{a}\, x^{\mu} = 0$$

Replacing definition (3.7), we obtain $\sum c_{a}a^{\mu}_{a} = 0$. Since the matrix $a^{i}_{a}(x)$ has rank equal to the smallest value between N and n, it follows that $c_{a} = 0$. Therefore, *the operators* X_{a} *are independent and generate a vector space called the* space of the linear operators of the Lie Group G, *denoted by* $\mathscr{G}$.

3.4.2 Infinitesimal Transformations on Vector Bundles

The situation here is similar to the previous case, with the difference that G acts on the fibers $\mathscr{V}$ of an arbitrary vector bundle, not necessarily resulting from a coordinate transformation.

Denoting a generic field defined in that vector bundle by $\Psi : \mathscr{M} \to T\mathscr{V}$ and denoting a field basis η_i, we may express the field as

$$\Psi = \Psi^i \eta_i$$

The infinitesimal transformation of the field is obtained as in the case of coordinates, where instead of a transformation of the coordinates x^μ we have a transformation of the components Ψ^i by the action of the group G, denoted by

$$\Psi'^i = f^i(\Psi^j, \theta^a)$$

which is followed by another transformation close to the first, given by

$$\Psi''^i = f^i(\Psi'^j, \theta^b + \delta\theta^b)$$

Expanding f^i in Taylor series around $\delta\theta^b = 0$ and keeping only the first powers of $\delta\theta^b$ we obtain as before

$$\Psi''^i = f^i(\Psi'^j, \theta^b) + \sum \frac{\partial f^i(\Psi'^j, \theta^b + \delta\theta^b)}{\partial\theta^c}\bigg\rfloor_{\delta\theta=0} \delta\theta^c = \Psi'^i + a^i_b(\Psi)\delta\theta^b \tag{3.8}$$

where we have denoted $a^i_b(\Psi) = \frac{\partial f^i}{\partial\theta^b}\Big\rfloor_{\delta\theta=0}$. Therefore, the infinitesimal variations of the field components are

$$\delta\Psi^i = a^i_b(\Psi)\delta\theta^b \tag{3.9}$$

and the infinitesimal variation of a function (or better, of a functional of the field such as, for example, the Lagrangian), of the field $F(\Psi)$, resulting from the above infinitesimal transformation is

$$\delta F = \frac{\partial F}{\partial\Psi^i}\delta\Psi^i = \frac{\partial F}{\partial\Psi^i} a^i_b(\Psi)\delta\theta^b = \delta\theta^b X_b F$$

where we have denoted the linear operators

$$X_a = \sum a^i_a(\Psi)\frac{\partial}{\partial\Psi^i} \tag{3.10}$$

These operators act on the space of all differentiable functions on $\mathscr{M}$.

In particular, applying X_a to the field components Ψ^α we obtain

$$X_a \Psi^i = a_b^i \tag{3.11}$$

The commutator or *Lie bracket* between these operators is defined by

$$[X_a, X_b]F = X_a(X_b F) - X_b(X_a F) \tag{3.12}$$

where F is an arbitrary function. Replacing the above expressions for X_a we find that

$$[X_a, X_b] = \left(a_a^k \frac{\partial a_b^j}{\partial \Psi^k} - a_b^k \frac{\partial a_a^j}{\partial \Psi^k} \right) \frac{\partial}{\partial \Psi^j} \tag{3.13}$$

A non-trivial result obtained by Marius Sophus Lie in 1872 shows that for a Lie group the commutators between two linear operators define an algebra in the space $\mathscr{G}$ generated by $\{X_\alpha\}$. The mathematical and physical implications of Lie's theorem reside in the fact that under the conditions defining the Lie group, it is sufficient to work with the above-mentioned algebra of linear operators.

Except for a few discrete symmetries, all relevant symmetries of the fundamental interactions satisfy these conditions. *The result of Lie is part of the development of field theory and particle physics from the 20th century onward. Actually, the results derived by Noether and Wigner suggest that any present or future theoretical proposals to modify the Lie symmetry structure must be checked against the theoretical and experimental results that are currently dependent on the Lie theorem.* This classic and non-trivial theorem can be shown in different ways. In the following we present some of its details [62–64].

Theorem 3.1 (Lie) *The commutator between two linear operators X_a of a Lie group is a linear combination of these operators*

$$[X_a, X_b] = f_{ab}{}^c X_c$$

where $f_{ab}{}^c$ are constants, called the structure constants of the group.

Consider a Lie group G with parameters θ, acting on the fiber $\mathscr{V}_p$ of a vector bundle

$$(\mathscr{M}, \pi, \mathscr{V})$$

Consider two consecutive infinitesimal transformations of the Lie group, one with increment $\theta + \delta\theta$ to the original values θ and the other with increments $\theta + d\theta$. From (3.1) and the definition of a Lie group, it follows that there is an analytic function between these parameters:

$$\theta^a + d\theta^a = \phi^a(\theta,\ \theta + \delta\theta)$$

The first-order terms of the Taylor expansion of this function around the identity (here represented by $\theta = 0$) gives a relation between the increments $d\theta$ and $\delta\theta$ as

$$d\theta^a = \phi(0,0) + \sum \frac{\partial \phi^a(\theta, \theta + \delta\theta)}{\partial \delta\theta^b} \rfloor_{\theta=0}\, \delta\theta^b = \sum_b \mathscr{F}_b{}^a \delta\theta^b \tag{3.14}$$

where we have denoted

$$\mathscr{F}_b{}^a(\delta\theta) = \frac{\partial \phi^a(\theta, \theta + \delta\theta)}{\partial \delta\theta^b}\Bigg\rfloor_{\theta=0}$$

We note that $\phi(0,0)$ is the relation between the parameter of the identity transformation ($\theta = 0$) and itself, so that $\phi(0,0) = 0$.

From the existence of the inverse element of a group, it follows that the relation (3.1) must also be invertible. That is, there exists the inverse matrix $\mathscr{F}^{-1b}{}_a(\theta)$ and the inverse relation is given by $\delta\theta^a = \mathscr{F}^{-1a}{}_b(\delta\theta)d\theta^b$.

Replacing this in the differential $d\Psi = \Psi'' - \Psi'$ given by (3.8) we obtain

$$d\Psi^i = \frac{\partial f^i(\Psi^j, \theta^b)}{\partial \theta^b} d\theta^b = \sum a^i_a \mathscr{F}^{-1a}{}_b(\theta) d\theta^b$$

from which we obtain

$$\frac{\partial f^i}{\partial \theta^b} = \sum a^i_a \mathscr{F}^{-1a}{}_b(\theta) \tag{3.15}$$

Now, since the transformation function f^i are analytic in θ they satisfy the Leibniz derivative rule, and the last expression gives

$$\frac{\partial a^i_c \mathscr{F}^{-1c}{}_b}{\partial \theta^a} = \frac{\partial a^i_c \mathscr{F}^{-1c}{}_a}{\partial \theta^b}$$

or equivalently

$$a^i_c \left(\frac{\partial \mathscr{F}^{-1c}{}_b}{\partial \theta^a} - \frac{\partial \mathscr{F}^{-1c}{}_a}{\partial \theta^b} \right) + \mathscr{F}^{-1c}{}_b \frac{\partial a^i_c}{\partial \theta^a} - \mathscr{F}^{-1c}{}_a \frac{\partial a^i_c}{\partial \theta^b} = 0 \tag{3.16}$$

However, from (3.8), a^i_c depends on θ^a only through Ψ^i, so that

$$\frac{\partial a^i_c}{\partial \theta^a} = \frac{\partial a^i_c}{\partial \Psi^j} \frac{\partial \Psi^j}{\partial \theta^a} = \frac{\partial a^i_c}{\partial \Psi^j} a^j_d \mathscr{F}^{-1d}{}_a$$

Consequently, (3.16) becomes

$$a^i_c\left(\frac{\partial \mathscr{F}^{-1c}{}_b}{\partial\theta^a}-\frac{\partial \mathscr{F}^{-1c}{}_a}{\partial\theta^b}\right)+\frac{\partial a^i_c}{\partial\Psi^j}a^j_d\mathscr{F}^{-1d}{}_a\mathscr{F}^{-1c}{}_b-\frac{\partial a^i_c}{\partial\Psi^j}a^j_d\mathscr{F}^{-1d}{}_b\mathscr{F}^{-1c}{}_a=0$$

Therefore, multiplication of this by $\mathscr{F}^a_m\mathscr{F}^b_n$ gives

$$\left(\frac{\partial \mathscr{F}^{-1c}{}_b}{\partial\theta^a}-\frac{\partial \mathscr{F}^{-1c}{}_a}{\partial\theta^b}\right)\mathscr{F}^a_m\mathscr{F}^b_n a^i_c=-a^j_d\frac{\partial a^i_c}{\partial\Psi^j}\left(\delta^d_m\delta^c_n-\delta^d_n\delta^c_m\right)=a^j_m\frac{\partial a^i_n}{\partial\Psi^j}-a^j_n\frac{\partial a^i_m}{\partial\Psi^j} \tag{3.17}$$

defining a function of the parameters θ by

$$\left(\frac{\partial \mathscr{F}^{-1c}{}_b}{\partial\theta^a}-\frac{\partial \mathscr{F}^{-1c}{}_a}{\partial\theta^b}\right)\mathscr{F}^a_m\mathscr{F}^b_n=f^c_{mn}(\theta) \tag{3.18}$$

or equivalently

$$\frac{\partial \mathscr{F}^{-1b}{}_a}{\partial\theta^c}-\frac{\partial \mathscr{F}^{-1b}{}_c}{\partial\theta^a}=f^b_{mn}(\theta)\mathscr{F}^{-1m}{}_c\mathscr{F}^{-1n}{}_a \tag{3.19}$$

(3.17) can be written in the more compact form

$$a^j_m\frac{\partial a^i_n}{\partial\Psi^j}-a^j_n\frac{\partial a^i_m}{\partial\Psi^j}=f^c_{mn}(\theta)a^i_c \tag{3.20}$$

To end the theorem we have to show that $f^c_{mn}(\theta)$ are constants. For this purpose, take the partial derivatives of the above expression with respect to θ^b, obtaining

$$\frac{\partial f^c_{mn}(\theta)}{\partial\theta^b}a^i_c+f^c_{mn}(\theta)\frac{\partial a^i_c}{\partial\Psi^k}\frac{\partial\Psi^k}{\partial\theta^b}=\frac{\partial}{\partial\Psi^k}\left(a^j_m\frac{\partial a^i_n}{\partial\Psi^j}-a^j_n\frac{\partial a^i_m}{\partial\Psi^j}\right)\frac{\partial\Psi^k}{\partial\theta^b}$$

or

$$a^i_c\frac{\partial f^c_{mn}(\theta)}{\partial\theta^b}=\left[\frac{\partial}{\partial\Psi^k}\left(a^j_m\frac{\partial a^i_n}{\partial\Psi^j}-a^j_n\frac{\partial a^i_m}{\partial\Psi^j}\right)-f^c_{mn}(\theta)\frac{\partial a^i_c}{\partial\Psi^k}\right]\frac{\partial\Psi^k}{\partial\theta^b}$$

Since $f^a_{bc}(\theta)$ depend only on θ, the derivative of (3.20) with respect to Ψ^k gives

$$\frac{\partial}{\partial\Psi^k}\left(a^i_n\frac{\partial a^j_m}{\partial\Psi^j}-a^j_m\frac{\partial a^i_m}{\partial\Psi^j}\right)=f^c_{mn}(\theta)\frac{\partial a^i_c}{\partial\Psi^k}$$

so that the right-hand side of the previous expression vanishes and consequently

$$a^i_c\frac{\partial f^c_{mn}(\theta)}{\partial\theta^b}=0$$

Since the matrix a^i_b has rank less than or equal to the smallest value between N and n, it follows that

$$\frac{\partial f^c_{mn}(\theta)}{\partial \theta^a} = 0$$

showing that $f^b_{mn}(\theta)$ are in fact constants. Replacing this result in (3.13) we obtain the Lie theorem

$$[X_a, X_b] = f^c_{ab} X_c \tag{3.21}$$

This implies that the space generated by $\{X_a\}$ defines an algebra with the product (the *Lie product*) defined by the commutator. The result is called *Lie algebra* of the group G denoted by $\mathscr{G}$. The constants f^c_{ab} are antisymmetric in the lower indices $f^c_{ab} = -f^c_{ba}$.

Another property of the Lie algebra is that it is non-associative. Instead of the associativity, it satisfy the *Jacobi identity*

$$[[X_c, X_a], X_b] + [[X_b, X_c], X_a] + [[X_a X_b], X_c] = 0$$

or in terms of the structure constants

$$f^p_{ca} f^n_{pb} + f^p_{bc} f^n_{pa} + f^p_{ab} f^n_{pc} = 0 \tag{3.22}$$

One interesting aspect of Lie's theorem is that almost everything can be done within the Lie algebra, including the representations of the Lie group [62]. This is a consequence of the analytical property which implies that it is possible to recover the full group starting from the Lie algebra.

Theorem 3.2 (The Inverse of Lie Theorem) *A finite transformation of a Lie group G can be obtained from the converging series of infinitesimal transformations generated by its Lie algebra $\mathscr{G}$.*

In fact, consider a set of constants f^c_{ab} satisfying (3.19) and (3.20). This means that there are functions $\mathscr{F}^{-1a}{}_b$ and a^i_a satisfying the equations

$$\begin{cases} \dfrac{\partial \mathscr{F}^{-1b}{}_a}{\partial \theta^c} - \dfrac{\partial \mathscr{F}^{-1b}{}_c}{\partial \theta^a} = f^b_{mn} \mathscr{F}_c^{-1m} \mathscr{F}_a^{-1n} \\ a^j_m \dfrac{\partial a^i_n}{\partial \Psi^j} - a^j_n \dfrac{\partial a^i_m}{\partial \Psi^j} = f^c_{mn} a^i_c \end{cases} \tag{3.23}$$

Replacing $\mathscr{F}_b^{-1a}$ from (3.15) and applying the initial condition (corresponding to the identity transformation)

$$\mathscr{F}^a_b(0) = \delta^a_b \tag{3.24}$$

we may solve in principle (3.23) for the functions $\mathscr{F}^{-1b}{}_a$. However, it is easier and more intuitive to change the parameterization of the group to a more convenient one, called the *vector parameterization* defined by

$$\theta^a = s^a \tau$$

whose geometrical interpretation is as follows: each set of values $(\theta^1, \ldots, \theta^N)$ defines a straight line with parameter τ in the space of parameters, passing through the origin and with direction s^a. Then, each transformation of the group corresponds to a point in such line. The identity transformation (conventionally described by $\theta^a = 0$) corresponds to origin $\tau = 0$.

In this new parameterization the operation of the group in a space $\mathscr{V}$ can be described by the *line operator* $S(\tau) = S(s^1\tau, \ldots, s^N, \tau)$ such that for each set of constant values of $s^1, \ldots, s^N$, the operator depends only on τ defined in the straight line. Then, the transformation of the field from $\Psi^i(0)$ to $\Psi^i(\tau)$ can be represented by

$$\Psi^i(\tau) = S(\tau)\Psi^i(0) \tag{3.25}$$

where the operator $S(\tau)$ still needs to be defined. For this, consider the variation of the field along the line

$$\frac{d\Psi^i(\tau)}{d\tau} = \frac{\partial\Psi^i}{d\theta^a}\frac{d\theta^a}{d\tau} = \frac{\partial\Psi^i}{\partial\theta^a}s^a = s^a\mathscr{F}^{-1b}{}_a(s,\tau)X_b\Psi^i$$

where in the last equal sign we have used (3.15). Consequently, the variation of Ψ^α can be expressed as $\partial\Psi^i/\partial\tau = \partial S/\partial\tau\Psi^i(0)$. The derivative of (3.25) compared with the above expression gives a differential equation for $S(\tau)$:

$$\frac{dS(\tau)}{d\tau} = s^a\mathscr{F}^{-1b}{}_a(s,\tau)X_bS(\tau)$$

This equation can be integrated with the boundary condition $S(0) = 1$ at $\tau = 0$, compatible with (3.24), obtaining

$$\left.\frac{dS(\tau)}{d\tau}\right]_{\tau=0} = s^aX_a$$

From the analytic dependence on the parameters it follows that $S(\tau)$ is also analytic in τ, so that it can be represented by a converging positive power series

$$S(\tau) = S(0) + \tau\left.\frac{dS}{d\tau}\right]_{\tau=0} + \cdots$$

or, using the above initial conditions,

$$S(\tau) = 1 + \tau s^a X_a + \cdots$$

Therefore, for each straight line defined by the parameters s^a, we may obtain the finite operation of the group. Then, for other finite group operations we only add the rotations of the straight line around the origin. This completes the theorem.

To obtain the finite transformations of a field Ψ we just apply the operator $S(\tau)$ to $\Psi^i(0)$, obtaining

$$\Psi^i(\tau) = \Psi^i(0) + \tau s^a X_a \Psi^i(0) + \cdots$$

From the above theorems it follows also that the existence of Lie subgroups implies the existence of Lie subalgebras, that can be characterized by the structure constants.

As an example, if H is a subgroup of a Lie group G with p parameters, then the commutator of two linear operators of H belongs to H:

$$[X_a X_b] = f_{ab}^c X_c, \qquad c = 1, \ldots, p,$$
$$f_{ab}^c = 0, \qquad c = p + 1, \ldots, N.$$

In particular, using the structure constants we may characterize Lie invariant subalgebras. From this we may define a simple Lie algebra (when it does not have proper invariant subalgebras) which corresponds to a simple group. A semi-simple Lie algebra (which does not have any Abelian invariant subalgebras) also corresponds to a semi-simple Lie group [63, 64] (these properties make an interesting exercise).

Definition 3.4 (Adjoint Representations of Lie Algebras) A representations of a Lie algebra $\mathscr{G}$ is a homomorphism

$$\mathscr{R} : \mathscr{G} \to \mathscr{G}'$$

where $\mathscr{G}'$ is an algebra of linear operators on a space V, such that

$$\mathscr{R}([X_a, X_b]) = [\mathscr{R}(X_a), \mathscr{R}(X_b)]$$

From the definition of Lie algebra it follows immediately that

$$[\mathscr{R}(X_a), \mathscr{R}(X_b)] = f_{ab}^c \mathscr{R}(X_c) \tag{3.26}$$

A consequence of the inverse theorem of Lie is that the representation of a Lie algebra induces the representation of the corresponding Lie group.

Similar to the representations of groups, the representations of Lie algebras are not unique, as they depend on the choice of the representation space, on their action of the algebra on that space, and finally on the choice of the basis of the representation space. Once the action of the algebra and the representation space is chosen, we may finally take a basis of that space and apply the Lie algebra operators $\mathscr{R}(X_a)$ on that basis:

$$\mathscr{R}(X_a)(\eta_i) = \sum \mathscr{R}(X_a)^j{}_i \eta_j \tag{3.27}$$

where $\mathscr{R}(X_a)^j{}_i$ denote the components of the representation matrix.

One particularly interesting representation of a Lie algebra is defined by the space of the Lie algebra itself, using the basis $\{X_a\}$, the same where the structure constants were defined. In this basis the algebra acts in the following way:

$$\tilde{\mathscr{G}}(X_a)X_b \stackrel{\text{def}}{=} [X_a, X_b] = f_{ab}^c X_c$$

where we have used a special notation $\tilde{\mathscr{G}}$ for this representation, which explicitly tells that the representation space is the space of the Lie algebra itself.

Comparing with (3.27), the matrix elements of the adjoint representation associated with the basis (taking $\eta_i = X_a$) are $\tilde{\mathscr{G}}(X_a)X_b = \tilde{\mathscr{G}}(X_a)^c{}_b X_c$. Therefore, it follows from (3.26) that the matrix elements of the adjoint representation are the structure constants of the Lie algebra

$$\tilde{\mathscr{G}}(X_a)^c{}_b = f_{ab}^c \tag{3.28}$$

In the adjoint representation all relevant group quantities are characterized by the structure constants.

Definition 3.5 (Casimir Operators) Given two operators $A = A^a X_a$ and $B = b^b X_b$ defined in the adjoint representation of a Lie algebra $\mathscr{G}$, we may define the product of the two operators consistently with the Lie algebra product as

$$\tilde{\mathscr{G}}(A)\tilde{\mathscr{G}}(B)X_c = [A, [B, X_c]] = A^a b^b [X_a, [X_b, X_c]] = A^a B^b f_{bc}^m f_{am}^n X_n$$

Since $\{X_a\}$ are linearly independent vectors, the above expression defines a matrix with components

$$(\tilde{\mathscr{G}}(A)\tilde{\mathscr{G}}(B))_c^n = A^a B^b f_{bc}^m f_{am}^n$$

whose trace is

$$tr(\tilde{\mathscr{G}}(A)\tilde{\mathscr{G}}(B)) = \sum_c (\tilde{\mathscr{G}}(A)\tilde{\mathscr{G}}(B))_c^c = A^a B^b f_{bn}^m f_{am}^n$$

This is a symmetric bilinear form which defines a scalar product in the Lie algebra space $<,>: \mathscr{G} \times \mathscr{G} \to I\!R$. It can be written as

$$< A, B >= g_{ab} A^a B^b, \text{ where } g_{ab} = f_{am}^n f_{bn}^m \tag{3.29}$$

This product is called the *Killing form* [65]. When the coefficients g_{ab} define an invertible matrix, we obtain a scalar product defining a metric geometry in the space of the Adjoint representation.

The *Casimir operators* of a Lie algebra are defined in the basis $\{X_a\}$ by

$$
\begin{aligned}
C^2 &= g_{ab} X^a X^b = f^n_{am} f^m_{bn} X^a X^b \\
C^3 &= f^p_{am} f^m_{bn} f^n_{cp} \\
&\vdots \\
C^k &= f^{m_k}_{a_1 m_1} f^{m_1}_{a_2 m_2} \cdots f^{m_{k-1}}_{a_k m_k} X^{a_1} X^{a_2} \cdots X^{a_k}
\end{aligned}
$$

They are *invariant operators* in the sense that they do not depend on the choice of basis in the Lie algebra.

The importance of the Casimir operators resides in the fact that the classification of the unitary irreducible representations of a Lie algebra (or of a Lie group) is given by the eigenvalues of these operators. In particular, Eugene Wigner showed that in the case of the Poincaré group, there are only two Casimir operators: The eigenvalues of the operator C^2 (the mass operator) acting on a Hilbert space gives the mass of the relativistic particles. On the other hand the eigenvalues of the operator C^3 (the spin operator) gives the spins of these particles [30]. This result provided a deep insight into the structure of the physical manifold.

The spectrum of the eigenvalues of the spin operator is discrete (formed by integers and semi-integers). On the other hand the spectrum of the mass operator is continuous, with isolated points (that is, not all real values appear).