Contents

List of Figures

List of Tables

Preface

In theoretical physics, routine use is made of many properties, such as recurrence relations and addition theorems, of the special functions of mathematical physics. These properties are for the most part classical, and their derivations are usually based on the methods of classical analysis. The purpose of this book is to show how these functions are also related to the theory of group representations and to derive their important properties from this theory. This approach elucidates the geometric background for the existence of the relations among the special functions. Moreover, the derivations may be more rationally motivated than are the usual complicated manipulations of power series, integral representations, and so on. I hope that the reader may find in this book reasonably simple derivations of many of the relations commonly used in theoretical physics for which the proofs may otherwise be somewhat unfamiliar.

In order that the book be fairly self-contained, approximately the first third delves into a preliminary discussion of such topics as Lie groups, group representations, and so on. The remaining chapters are devoted to various groups, and the special functions are discussed in conjunction with the group with which it is associated. Because of the inclusion of the introductory material, the only prerequisite is a reasonable knowledge of linear algebra.

The original impetus for the writing of this book was provided by a lecture course given by Professor Eugene P. Wigner a number of years ago. I am greatly indebted to Professor Wigner for his suggestion that I pursue the subject of the lectures further and for his continued friendly interest and advice in the work. I wish to thank Dr. Trevor Luke for carefully checking the manuscript. I also wish to thank my wife, whose encouragement contributed greatly to the writing of the book.

James D. Talman

London, Canada July 1968

General introduction

All of us have admired, at one time or another, the theory of the higher transcendental functions, also called special functions of mathematical physics. The variety of the properties of these functions, which can be expressed in terms of differential equations which they satisfy, in terms of addition theorems or definite integrals over the products of these functions, is truly surprising. It is surpassed only by the variety of the properties of the elementary transcendentals, that is the exponential function, and functions derived therefrom, such as the trigonometric functions. At the same time, special functions, as their full name already indicates, appear again and again as solutions of problems in theoretical physics.

These higher transcendentals are analytic functions of their arguments and their properties are usually derived on the basis of their analytic character, using the methods of the theory of analytic functions. Neither the present volume, nor the earlier lectures which gave the incentive to this volume, intend to question the beauty of the analytic theory of the special functions, nor the generality of the results which this theory furnishes. In fact, the lectures started with the observation that the results of the analytic theory are more general than those furnished by the method to be employed in the lectures, thus pointing to a drawback of the considerations to be presented. Though this has been substantially eliminated by subsequent developments, presented also in the present volume and at least partially in other publications, this in no way diminishes the beauty and elegance of the analytic theory, or the inventiveness that was necessary to its development.

Rather, the claim of the present volume is to point to a role of the "special functions" which is common to all, and which leads to a point of view which permits the classification of their properties in a uniform fashion. The role which is common to all the special functions is to be matrix elements of representations of the simplest Lie groups, such as the group of rotations in three-space, or the Euclidean group of the plane. The arguments of the functions are suitably chosen group parameters. The addition theorems of the functions then just express the multiplication laws of the group elements. The differential equations which they obey can be obtained either as limiting cases of the addition theorems or as expressions of the fact that multiplication of a group element with an element in the close neighborhood of the unit element furnishes a group element whose parameters are in close proximity of the parameters of the element multiplied. The integral relationships derive from Frobenius' orthogonality relations for matrix elements of irreducible representations as generalized for Lie groups by means of Hurwitz's invariant integral. The completeness relations have a similar origin. Further relations derive from the possibility of giving different equivalent forms to the same representation by postulating that the representatives of one or another subgroup be in the reduced form. Finally, some of the Lie groups can be considered as limiting cases of others; this furnishes further relations between them. Thus, the Euclidean group of the plane can be obtained as a limit of the group of rotations in three-space. Hence, the elements of the representations of the former group (Bessel functions) are limits of the representations of the latter group (Jacobi functions).

Because of the important role which representations of simple groups play in problems of physics, the significance of the “special functions” in physical theory also becomes more understandable. In fact, it appears that the elementary transcendentals are also “special functions,” the corresponding group being the simplest Lie group of all: the one parametric and hence Abelian Lie group. On the other hand, the field opens up in the opposite direction and one will wonder, when reading this book, what the role and what the properties are of representation coefficients of somewhat more complex Lie groups. Naturally, the common point of view from which the special functions are here considered, and also the natural classification of their properties, destroys some of the mystique which has surrounded, and still surrounds, these functions. Whether this is a loss or a gain remains for the reader to decide.

Eugene P. Wigner

Chapter 1

Introduction

In addition to the elementary transcendental functions such as e^x and $\sin x$, an important role is played in mathematical physics by the *special functions.* Examples of these functions are the Bessel functions, Legendre functions, and hypergeometric functions. For the most part the properties of these functions are studied on the basis of their analytic properties as solutions of ordinary differential equations. For example, the Bessel functions In are the solutions of the differential equation

$$J_n''(x) + \frac{1}{x} J_n'(x) + \left(1 - \frac{n^2}{x^2}\right) J_n(x) = 0 \tag{1.1}$$

that behave as $(x_n) / (2^n n!)$ for $x \to 0$. The Bessel functions are analytic functions of their argument and their order n, although in the original form of (1.1) the variable x was taken to be real, and the order is frequently restricted to be an integer.

The special functions are treated from this point of view in many excellent books. The best known among physicists are those of Courant and Hilbert [1] and Morse and Feshbach [2] but there are many other treatises, such as those of Rainville [3] and Lebedev [4] devoted to the subject.

The purpose of the present monograph is to demonstrate some of the properties of special functions from the point of view of group theory, or more specifically, from the theory of group representations. It will be seen that many of the special functions are matrix elements, or are simply related to matrix elements, of the representations of elementary groups such as rotation groups and Euclidean groups. Many properties of the special functions can then be derived from a unified point of view from the group representation property. For example, the Legendre functions are matrix elements of representations of the rotation group in three dimensions. The addition theorems for these functions then follow from the group multiplication law. The differential equations for Legendre functions are a consequence of the differential equations that relate the derivatives of group representations to the corresponding representations of the Lie algebra of the group. The orthogonality and completeness relations are the orthogonality and completeness relations of the group representations. Further relations can be obtained by transforming a given representation to an equivalent form and by reducing the direct product of two representations into a sum of irreducible representations.

The group theoretic treatment shows that the special functions are special only in that they are related to specific groups. The usefulness of group representation theory for the solution of a variety of physical problems makes it natural that representation matrix elements are important special functions for many problems in mathematical physics. It may further be true that the properties that can be derived group theoretically are their most important ones, since they originate from the "geometric" properties of the functions.

Although it provides a unified basis for the treatment of special functions, the group theoretic approach has a number of limitations. Not all special functions arise as elements of group representation matrices; for example, no group theoretic basis is known for the gamma and elliptic functions. The special functions that occur in group representations have restricted indices; for example, only the Legendre functions of the first kind of integer order arise in a natural way. Certain other properties, such as the many integral representations, are not obvious consequences of the representation property. The special functions that will be considered in detail in this work are the complex exponential function, Jacobi functions and Legendre functions (which are related to hypergeometric functions), Bessel and spherical Bessel functions, Gegenbauer polynomials, associated Laguerre polynomials, and Hermite polynomials. These arise in connection with the groups of pure rotations in two, three, and four dimensions, the Euclidean groups (rigid transformations) in two and three dimensions, and a less familiar group that corresponds to the Lie algebra generated by the position and momentum operators of quantum mechanics.

The approach that we will follow has a certain resemblance to one that has received considerable attention recently and that is related to the *factorization method* of Infeld and Hull [5]. In the factorization method, a single second-order differential equation is replaced, if possible, by a pair of first-order differential equations for a whole set of special functions, that is, a pair of equations of the form

$$L_n^+ f_n = f_{n+1}, \qquad L_n^- f_n = f_{n-1} \tag{1.2}$$

where L_n^+ and L_n^- are first-order differential operators. The second-order equation can then be written in the two alternate forms

$$L_{n+1}^- L_n^+ f_n = f_n, \qquad L_{n-1}^+ L_n^- f_n = f_n. \tag{1.3}$$

It is possible to identify the operators L_n^+ and L_n^- (together with additional operators) with a Lie algebra, and the possible factorizations can be classified by the study of these Lie algebras. The special functions constitute basis functions in representation spaces (as will be defined), for Lie algebras and many of their properties can be obtained in this way. This approach has been thoroughly investigated by W. Miller [6] and B. Kaufman [7].

The approach that will be followed here differs from this in that the primary emphasis will be on groups rather than on the corresponding Lie algebras, and most of the special functions will be related to matrix elements of group representations rather than basis functions in representation spaces, although some interesting results will also be obtained in this way.

A large amount of work has been done in the past few years on the relationship of special functions to group representations, particularly by Miller [8], and a book by N. I. Vilenkin [9] on the subject has been published in Russian. Our attention in this book will be limited for the most part to those properties that seem to be of most interest for mathematical physics. A considerable part of the material in this book arises from the lecture notes of Wigner on the subject [10].

Chapter 2

Abstract groups

It seems natural to commence the study of the application of group theory to the special functions with a review of group theory itself. This chapter is included to meet this possible need and to define some of the terms that are met in the remainder of the book. The contents of this chapter will be familiar, at least in outline, to many readers, and they are .invited either to omit it or to take pleasure in the review of familiar concepts. This chapter is for this reason brief and rather tedious, without the examples and observations which give rise to the interest in the subject.

It is instructive, in the study of group theory, to view it as an investigation of the $1{-}1$ mappings of a set S onto itself. The group elements are functions f defined for all $x \in S$ such that $f(x) \in S$, that is, the domain and range of f are both S. The function f must be $1{-}1$; that is, for each $y \in S$, there must be a unique x such that $y = f(x)$. A group composed of such mappings is called a *transformation group*. As an example, the set S might consist of the first n integers, in which case the group is the permutation group on n objects.

The requirement that the mapping be $1{-}1$ implies that it has a well-defined inverse f^{-1}; f^{-1} is the mapping that undoes the operation of f.

If f and g are $1{-}1$ mappings, the composition mapping gf defined by $[gf](x) = g(f(x))$ is easily seen to be $1{-}1$. This mapping is called the product of g and f. It is this type of multiplication that is the basic operation in group theory, adjoining an element, the product gf, to each ordered pair of elements g and f.

Some groups are sets of mappings of a set onto itself that are restricted in some further way than the $1{-}1$ condition that has been imposed. For example, the group of transformations of a plane into itself that keep the distance between points fixed and also keep one point fixed is the rotation group in the plane. It is with groups of this nature that we will be principally concerned.

2.1 Abstract groups

A group G is a set in which an operation is defined which associates with every ordered pair of elements in G a third element of G. This operation is called *multiplication*; each of the given pair is called a *factor*, and the third element is called the *product*. If (a, b) is the given pair, the product is usually denoted simply by ab. The set G and the multiplication law must further satisfy the following properties known as group axioms.

A. The multiplication is *associative*; that is

$$(ab)\,c = a\,(bc)\,. \tag{2.1}$$

B. There is one element, *the identity* e, with the property that for all $a \in G$

$$ae = ea = a. \tag{2.2}$$

C. For each $a \in G$, there is an element a^{-1}, the *inverse* of a, such that

$$a^{-1}a = aa^{-1} = e. \tag{2.3}$$

An important consequence of A is that the product of three (or more) factors in a particular order is independent of the order in which the multiplications are performed. Thus, a product of the form $abcd$ can be interpreted to be anyone of $(ab)(cd)$, $(a(bc)d)$, $((ab)c)d$, and so on. It is important to observe that either of the equations $ax = b$ and $xa = b$ has a unique solution for x. In the former case $x = a^{-1}b$, in the latter $x = ba^{-1}$, as can be seen by left (right) multiplication with a^{-1}. It follows from this that if $ab = b$, or if $ba = b$, then $a = e$, since $a = e$ is a solution of either equation. Therefore, e is the only element multiplication with which leaves any element unchanged, and this applies to both left and right multiplication. Similarly, if $ab = e$, or $ba = e$, then $b = a^{-1}$. The inverse element is therefore unique; that is, $b = a^{-1}$ is the only element with the property C.

If $a^2 = a$, multiplication with a^{-1} shows that $a = e$: the identity is the only element equal to its square (although it is possible that $a^n = a$, $n \neq 2$). Since $e^2 = e$, the identity element is its own inverse.

Another rule which is frequently used is that the inverse of the product ab is $b^{-1}a^{-1}$.

A result of considerable significance is that the mapping f_a of G into G defined by $f_a\left(a^{-1}y\right) = aa^{-1}y = y$ is a $1{-}1$ correspondence between the group and itself. For any $y \in G$, $f_a\left(x_1\right) = f_a\left(x_2\right)$, so that y is the image of an element in G, that is, f_a maps G onto G. Furthermore, if $f_a\left(x_1\right) = f_a\left(x_2\right)$, $x_1 = x_2$, since multiplying each side of $ax_1 = ax_1$ with a^{-1} gives $x_1 = x_2$. Therefore, each $y \in G$ is the image of exactly one $x \in G$ and the mapping f_a has a well-defined inverse. The mapping f_a in effect rearranges or permutes the elements of G.

Each group element can therefore be identified with a unique map- ping of the group onto itself with the property that the product $x_2^{-1}x_1$ is invariant. It is obvious that each group element generates such a mapping. Conversely, if F is such a mapping, the group element is such that $F = f_a$ can be identified as $a = F(x)x^{-1}$; the invariance of $x_2^{-1}x_1$ guarantees that this choice is independent of x.

We consider now some examples of groups. One simple example is the set C of all nonzero complex numbers under complex multiplication. Another example is the set of all nonsingular $n \times n$ matrices with complex elements; the multiplication in this group, and in all matrix groups to be considered in this work, is the usual matrix multiplication.

A third example has been alluded to in the introduction. It is the *permutation group* which can be defined as follows. For any set S the set of all mappings f of S onto itself which have a well-defined inverse constitutes a group with the multiplication defined as in the introduction. The identity of the group is the mapping defined by $f(x) = x$. It is necessary to verify that the associative law is valid, that is, that $(fg)h = f(gh)$; this can be seen to be true since each of these is the mapping $x \longrightarrow f(g(h(x)))$.

It may happen that two groups, which are defined in quite different ways, are identical as far as their mathematical structure is concerned. If this is the case, the groups are said to be isomorphic. The precise definition of isomorphism is as follows: two groups G and G' are *isomorphic* if there is a $1{-}1$ correspondence between them such that if $a \longleftrightarrow a'$ and $b \longleftrightarrow b'$, then $ab \longleftrightarrow a'b'$ for all a and b in G. This correspondence is called an *isomorphism.* One can see immediately that $e \in G$ must correspond to e', the identity in G', by the following argument. Suppose $e \longleftrightarrow c'$. Then for any a and corresponding a', $ae \longleftrightarrow a'c'$ and, hence, $a \longleftrightarrow a'c'$. This implies that $a'c' = a'$ and, hence, as we have seen, that $c' = e'$. It can also be proved easily that if $a \longleftrightarrow a'$, $a^{-1} \longleftrightarrow (a')^{-1}$. As a simple example of isomorphism, we remark that the group of complex numbers mentioned above is isomorphic to the group of matrices of the form

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

where a and b are real and not both zero and the above matrix corresponds to the complex number $a + bi$ under the isomorphism.

If two elements have the property that $ab = ba$, they are said to *commute.* A group that has the property that all pairs of elements commute is said to be *Abelian*, or *commutative.* One can devise many examples of Abelian groups; the simplest is, perhaps, the set of real numbers with addition as the group operation. We mention also that if a group has a finite number n of elements, it is said to be a finite group of order n.

2.2 Subgroup and factor groups

A subset of a group G which is itself a group with the same law of multiplication as that in G is called a *subgroup.* To be more explicit, a subgroup H is a nonempty subset of G such that if a and b are in H, then a^{-1} and ab are in H. It is unnecessary to specify the associative property since this is guaranteed by the multiplication law in G. Furthermore, $e \in H$ since if $a \in H$, $a^{-1} \in H$, and $aa^{-1} = e \in H$. Two trivial examples of subgroups of G are the group G itself and the subgroup consisting of only e. If a subgroup is neither of these, it is said to be a *proper* subgroup. An example of a subgroup is the set of all $n \times n$ matrices with unit determinant as a subgroup of the set of all $n \times n$ matrices. Another example is provided by the group of permutations on a set S. If S' is a subset of S, the set of permutations that have the property $f(x) = x$ for all $x \in S'$ is a subgroup of the full permutation group; this subgroup is in fact the permutation group on the set $S - S'$. Another subgroup, which is larger, is the set of all permutations that map S onto itself and $S - S'$ onto itself.

If a is any element of a group G and B is any subset of the group, it is convenient to denote by aB the set of all elements of the form ab where $b \in B$; Similarly Ba denotes the set of all elements of the form ba. If A and B are subsets, we will denote by AB and BA respectively the sets of all elements of the form ab and ba where $a \in A$ and $b \in B$. These sets satisfy certain associativity properties such as $(AB)C = A(BC)$, which are easily verified and will henceforth be taken for granted. As examples of this notation, we observe that $eA = A$ for any A, and that if H is a subgroup, $HH = H^2 = H$. We will also denote by A^{-1} the set of all inverses of elements of A. We can observe that $(AB)^{-1} = B^{-1}A^{-1}$ and that if H is a subgroup, $H^{-1} = H$. It is also true that if $H^2 = H^{-1} = H$, H is a subgroup.

A subgroup H of a group G can be used to subdivide G into disjoint pieces known as the *left cosets* of H. These are the subsets of G of the form aH where a is any element of G. Each element x of G is in exactly one left coset. It is certainly in the left coset xH since $x = xe \in xH$. Suppose, however, that x is in two cosets, aH and bH. Then x can be written either as ah_1 or as bh_2,where h_1 and h_2 are elements of H. Then $ah_l = bh_2(= x)$ or $a = bh_2h_1^{-1}$ and $aH = bh_2h_1^{-1}H = bh_2H = bH$ so that aH and bH are identical. We remark further that a and b are in the same left coset if and only if a can be expressed in the form bh, $h \in H$; this is equivalent to the condition $b^{-1}a \in H$, which is a useful criterion for determining whether two elements are in the same left coset.

If the subgroup H contains a finite number m of elements, and there are p, a finite number, left cosets, the group G contains $p \times m$ elements since each left coset contains exactly m elements. This implies that in a finite group the order of a subgroup must divide the order of the group, a result known as Lagrange's theorem.

It is clear that *right cosets* can be defined analogously to left cosets and that the foregoing remarks will be equally applicable. We observe that two elements a and b are in the same right coset if, and only if, $ba^{-1} \in H$. In general the right cosets will differ from the left cosets; if, however, $aH = Ha$ for all $a \in G$, the two types of cosets coincide and the subgroup H is said to be *normal* or *invariant.* It can be seen that a condition equivalent to $aH = Ha$ for all $a \in G$ is that $aHa^{-1} \subset H$ for all $a \in G$, and the condition of normality will usually be expressed in this form. If a group has no proper normal subgroup, it is called *simple.*

If N is a normal subgroup of G, an important new group can be formed which is known as the *factor* or *quotient* group and is denoted by G/N. The elements of G/N are the cosets of N. Multiplication is defined in G/N by defining the product of two cosets aN and bN to be the coset

$$(aN)(bN) = aNbN = abNN = abN;$$

that is, the product of aN and bN is the coset containing ab. (It is not difficult to show that this definition is independent of the choice of a and b from their respective coset.) The identity of G/N is N itself since

$$(aN)N = aNN = aN.$$

The associative law is readily seen to be valid since

$$(aNbN)\, cN = abcN = aN\, (bNcN)\,.$$

Finally, the inverse of aN is clearly $a^{-1}N$.

An example of a normal subgroup is the group of matrices of unit determinant as a subgroup of a complete matrix group; it will be left to the reader to show that this is a normal subgroup, and that the corresponding factor group is isomorphic to the group of nonzero complex numbers.

A few rather elementary properties of subgroups will now be formulated as a theorem, for which the proofs will not be given.

Theorem 2.1. *Let G be a group.*
(a) If H and K are subgroups of G, then $H \cap K$, the set of elements contained in both H and K, is a subgroup of G.
(b) If H is a subgroup of G, and N is a normal subgroup of G, then $N \cap H$ is a normal subgroup of H.
(c) If H is a subgroup of G, and N is a normal subgroup of G, then NH is a subgroup of G.
(d) If H and N are normal subgroups of G, then NH and $H \cap N$ are normal subgroups of G.

2.3 Homomorphisms

A more general relation between two groups than isomorphism is that of homomorphism. A mapping f from a group G onto a group G' is said to be a *homomorphism* if for every pair of elements a, b in G, $f(a)f(b) = f(ab)$. (Note that $f(a)$, $f(b)$, and $f(ab)$ are elements of G'.) A homomorphism is more general than an isomorphism in that several elements of G may be mapped into a single element of G', so that f may not have a single-valued inverse. It is required, however, that every element of G' be the image of some element of G. (This is implied by the preposition "onto.")

In this section a few simple properties of homomorphisms will be described. The image of e, the identity in G, is e', the identity in G'; that is, $f(e) = e'$. This can be shown in the way that the same result was proved for isomorphisms. Similarly, the images of reciprocals are again reciprocals: $f(a)f(a^{-1}) = f(e) = e'$ or $f(a^{-1}) = (f(a))^{-1}$.

It is interesting to observe that the set of elements of G that are mapped by f into the identity e' of g' is a normal subgroup of G. This set is called the *kernel* of f and is often denoted by K. An element k of G is in K if $f(k) = e'$. If k_1 and k_2 are both in K, $f(k_1k_2) = f(k1)f(k2) = e, 2 = e'$ showing that k_1k_2 is also in K. Moreover, if $f(k) = e'$, $f(k^{-l}) = e'^{-1} = e'$, showing that $k^{-1} \in K$ and hence that K is a subgroup. To show that K is a normal subgroup, we consider its cosets. It will be shown that all elements of a coset of K map onto a single element of G', and conversely, that all the elements of G that are mapped onto a single element of G' belong to the same coset of K. Let aK be a left coset of K and suppose $x \in aK$. Then $x = ak$, $k \in K$, and $f(x) = f(a)f(k) = f(a)$, showing that all elements of aK are mapped by f into $f(a)$. Conversely, if $f(x) = f(a)$, $f(a^{-1}x) = f(a)^{-1}f(x) = e'$, implying that $a^{-1}x \in K$ and that $x \in aK$. It can be concluded that the coset aK is the set of all elements x such that $f(x) = f(a)$. These considerations apply just as well to Ka so that $aK = Ka$ and K is normal.

It is evident that there is a 1-1 correspondence between cosets of K and elements of G' defined by the relation

$$aK \longleftrightarrow f(a) \tag{2.4}$$

Furthermore, this $1-1$ correspondence is an isomorphism between G' and the factor group G/K composed of the cosets of K, as is shown by the calculation $(aK)(bK) = abK$?? $f(ab) = f(a)f(b)$. These findings can be summarized as a theorem.

Theorem 2.2. *Let f be a homomorphic mapping of a group G onto a group G'. Then the set K of elements of G that map onto the identity of G' is a normal subgroup of G, the set of elements of G that map onto an element a' of G' is a coset of K, and G/K is isomorphic to G'.*

If N is a normal subgroup of G, it is easy to see that the mapping of G onto G/N that maps each element of G onto the coset of which it is a member is a homomorphism. This is known as the natural homomorphism *of G onto G/N; the kernel of the natural homomorphism is obviously N, since N is the identity of G/N.*

2.4 Some further aspects of groups

Our review of some of the concepts from abstract group theory will be concluded in this section with discussions of miscellaneous topics: equivalence classes in groups, direct products of groups, and two important subgroups. The discussion will be limited to definitions and results more or less pertinent to the sequel.

An important classification of elements within a group G is provided by the conjugacy or equivalence classes in the group defined as follows. An element $b \in G$ is said to be *equivalent*, or *conjugate*, to $a \in G$ if there is some element t such that $tat^{-1} = b$. If this is the case, t is said to *transform* a into b. Since e transforms a into a, a is equivalent to itself, and since t^{-1} transforms b into a, a is equivalent to b. Furthermore, if c is equivalent to b, that is, $c = sbs^{-1}$, and b is equivalent to a, $b = tat^{-1}$, then c is equivalent to a since $c = (st)a(st)^{-1}$. The group G can be divided into disjoint subsets with the property that any two elements common to one subset are equivalent and any two elements from different subsets are not equivalent; these subsets are called the *equivalence* or *conjugacy classes* of G. The class to which any element a belongs will be denoted by C_a. As an example, we note that C_e consists only of e since $tet^{-1} = e$ for all $t \in G$. If G is Abelian, it is clear that each class consists of only a single element.

It is worth while to observe that a subgroup N of G is normal if and only if it is composed of entire classes. This follows since, if $a \in N$,$tat^{-1} \in N$ if N is normal, so that $C_a \subset N$. On the other hand, if $C_a \subset N$ for each $a \in N$, $tNt^{-1} \subset N$ and N is normal. The subgroup N can itself be regarded as a group composed of classes $\overline{C}$ with respect to N. Two elements n_1 and $n2$ are in the same class C if there is an element $t \in N$ such that$tn_1t^{-t} = n_2$; it follows that if two elements are in the same class $\overline{C}$ of N, they are in the same class C of G. On the other hand, two elements common to a class C in G may be in different classes in N, since the element which transforms one to the other may be outside of N. We can conclude from this discussion that each class C in G that is contained in N is composed of entire classes $\overline{C}$ of N.

If two groups G_1 and G_2 are known, it is possible to form a third group from them. The *direct product* $G_1 \otimes G_2$ of the two groups is defined to be the set of all pairs of the form (g_1, g_2) where $g_1 \in G_1$ and $g_2 \in G_2$ The product of two such pairs is defined in a rather obvious way by

$$(g_1 g_2)(h_1, h_2) = (g_1 h_1, g_2 h_2). \tag{2.5}$$

The identity in $G_1 \otimes G_2$ is (e_1, e_2) and the inverse of (g_1, g_2) is $\left(g_1^{-1}, g_2^{-1}\right)$.

The direct product has no interesting algebraic structure other than that of each of its factors. A question of some interest is: given a group G, does it have subgroups G_1 and G_2 such that G is (isomorphic to) the direct product of G_1 and G_2? It will now be shown that if G has normal subgroups N_1 and N_2 such that $N_1 \cap N_2 = \{e\}$ and $N_1 N_2 = G$, then G is isomorphic to $N_1 \otimes N_2$. It will first be shown that if $n_1 \in N_1$ and $n_2 \in N_2$, then $n_1 n_2 = n_2 n_1$. This is proved by considering the **commutator** $q = n_1 n_2 n_1^{-1} n_2^{-1}$ of n_1 and n_2. Since N_2 is a normal subgroup, $n_1 n_2 n_1^{-1} = n_2' \in N_2$ and $q = n_2' n_2^{-1} \in N_2$. Similarly, since N_1 is a normal subgroup, $n_2 n_1^{-1} n_2^{-1} = n_1' \in Nt$ and $q = n_1 n_1' \in N_1$. Therefore, $q \in N_1 \cap N_2$ and q is necessarily e. It follows immediately that $n_1 n_2 = n_2 n_1$. Since $N_1 N_2 = G$, each element $g \in G$ can be expressed in the form $n_1 n_2$. This representation is, moreover, unique, since if $g = n_1 n_2 = m_1 m_2$ where n_1 and m_1 are in N_1 and n_2 and m_2 are in N_2, then $m_1^{-1} n_1 = m_2 n_2^{-1}$. The left-hand side is in N^{-1} the right-hand side is in N_2, so that each is necessarily e and $m_1 = n_1$, $m_2 = n_2'$. The isomorphism between $N_1 \otimes N_2$ and G is now given by the relation

$$(n_1, n_2) \longleftrightarrow n_1 n_2 \tag{2.6}$$

This correspondence is an isomorphism, since if $(n_1, n_2) \longleftrightarrow n_1 n_2$ and $(m_1, m_2) \longleftrightarrow m_1 m_2$, then

$$(n_1, n_2)(m_1, m_2) = (n_1 m_1, n_2 m_2) \longleftrightarrow (n_1 m_1 n_2 m_2) = (n_1 n_2)(m_1 m_2)$$

from the fact that $n_2 m_1 = m_1 n_2$.

The set Z of all elements that commute with every element of a group G is called the **center** of G. It is easy to show that Z is a subgroup of G, since if z_1 and z_2 are in Z and g is any element of G,

$$z_1 z_2 g = z_1 g z_2 = g z_1 z_2$$

and $z_1 z_2 \in Z$. It is also clear that $e \in Z$ and that if $z \in Z$, $z^{-1} \in Z$, so that Z is a subgroup. Since $gZ = Zg$ for all $g \in G$, Z is also a normal subgroup of G. Any subgroup of Z is also necessarily a normal subgroup of G; such a subgroup is called a **central normal subgroup**.

We have defined the commutator of two elements a and b to be $aba^{-1}b^{-1}$. The **commutator subgroup** of G is defined to be the set of all elements that can be expressed as a product of commutators. The commutator subgroup is commonly denoted by G'. It is clear that the product of two elements of G' is again in G' and that $e \in G'$. If q is the commutator of a and b, q^{-1} is the commutator of b and a. It follows from this that the inverse of any element in G' is again in G', and, hence, that G' is a subgroup. To show that G' is a normal subgroup, we observe first that if q is the commutator of a and b, tqt^{-1} is the commutator of tat^{-1} and tbt^{-1} so that t transforms any commutator into another commutator. If $x = q_1 q_2 \dots q_n$ is an element of G', then

$$txt^{-1} = \left(tq_1t^{-1}\right)\left(tq_2t^{-1}\right)\dots\left(tq_nt^{-1}\right) \tag{2.7}$$

is also an element of G', which is, therefore, a normal subgroup.

An interesting property of G' is that G/G' is an Abelian group. Let aG' and bG' be two cosets of G'. Their inverses in G/G' can be written $a^{-1}G$, and $b^{-1}G'$ respectively. The commutator of aG' and bG' is

$$(aG')(bG')\left(a^{-1}G'\right)\left(b^{-1}G'\right) = aba^{-1}b^{-1}G' = G',$$

since $\left(a^{-1}G'\right)\left(b^{-1}G'\right) \in G'$. Since G' is the identity in G/G', G/G' is Abelian.

Chapter 3

Lie Groups

Although the study of group theory originated with finite groups, the theory of Lie groups has achieved greater importance than that of finite groups, and it is Lie groups that are of primary significance for our purposes. Whereas the elements of a finite group form a discrete set, those of a Lie group form a continuum; they depend differentiably on one or more parameters. Because of this, topological considerations, the concepts of neighborhood, open sets, closed sets, and so on, play an important role in the basic theory of Lie groups. Many remarkable results, for example that continuity of the group implies differentiability of the group under certain circumstances, have been established for Lie groups. However, in the present discussion we will avoid involved topological considerations as much as possible, often by making more restrictive assumptions than are necessary (differentiability instead of continuity, and so on).

Many of the properties of a Lie group are determined by the elements in an arbitrarily small neighborhood of the identity, since such a neighborhood can be multiplied by itself repeatedly to generate at least a large part of the group. For example, if the elements in any neighborhood of the identity commute, the group is Abelian. There is no corresponding situation in finite groups, in which a small subset cannot be expected to determine most of the group properties.

Whereas the theory of finite groups can stand alone, without relying on any other part of mathematics, the theory of Lie groups often makes extensive use of the theory of ordinary and partial differential equations. In addition, many Lie groups have connectedness properties which, if perhaps not complicated, are still different from the connectedness properties of ordinary Euclidean spaces. An example of this, that of the group of rotations in two dimensions, will be found in the next section. There are Lie groups with even more complicated "topology in the large." As a result of these circumstances, particularly the reliance on the theory of differential equations, the theory of Lie groups is much less independent of other parts of mathematics than is the theory of finite groups.

3.1 Lie Groups

A *Lie group* of dimension n is a group with the following properties.

1. The group is composed of a finite number of subsets, which are not necessarily disjoint, such that the elements of each subset can be placed in 1–1 correspondence with the points of open regions in n–dimensional space. These regions will be called *parameter domains* and the n coordinates of a point in one of them will be called *group parameters* or *coordinates* and will be used as coordinates for the corresponding group element.

2. A given group element may have parameters in more than one parameter domain; if this is the case the coordinates in one region must be arbitrarily differentiable functions of the coordinates in the other region.

3. The group operations must be differentiable in the group parameters; the exact meaning of this condition will be explained shortly.

4. The group is closed (or has the closure property) and is connected. These properties will be defined precisely later.

It would be most convenient if one open region in the parameter space were sufficient to label all the group elements. Unfortunately, it is usually necessary for topological reasons, that is, reasons of connectedness, to introduce different coordinate systems into different parts of the group. For the most part, this complication will, however, be ignored and we may refer to the coordinates of a group element without mentioning explicitly that they are in a particular parameter domain.

The various parameter domains could be made to be nonoverlapping by discarding from some of them those elements that they have in common with others. The reason that this is not done is that the resulting domains, which would be differences of open sets, would not be open, and it is desirable that the parameter domains be open sets. In fact, one of the difference sets could conceivably he a single point, in which case the differentiability condition would become meaningless.

As an example of all this, consider the set of rotations of a plane, which can be labelled by the rotation angle θ.

This θ is, however, a cyclic variable and rotations by θ and by $\theta \pm 2\pi n$ (n an integer) are identical. It is, therefore, impossible to introduce a variable with a finite domain which would depend continuously on the group element: the group is, like the circle, multiply connected whereas any finite domain of a variable is singly connected. Hence, one introduces two domains to cover the group. The variable of the first domain is $\theta_1 = \theta$ and is restricted to the open region $-\pi < \theta_1 < \pi$. The second one is $\theta_2 = -\pi$ and can also be restricted to the domain $-\pi < \theta_2 < \pi$. The only point the first domain does not cover is the point $\theta = \pi$, the only point the second domain does not cover is $\theta = \pi$, that is, the unit element. Of course, both domains could be chosen to be smaller as long as they cover all points $-\pi < \theta \leq \pi$.

The reason for the need of several open domains in parameter space (actually only two for the groups to be considered) is always the same in principle, but the situation may be somewhat more complicated than in the case just considered.

A particular group will naturally admit many different coordinate systems, each of which will be called a *parametrization.*

The set of coordinates of a group element will be denoted by a boldface letter. We will frequently not distinguish between a group element and the point in the parameter space that represents it; for example, we may refer to the product of elements **p** and **q**.

It is convenient to assume that the identity element has coordinates zero in at least one parameter domain. This is no real restriction since it can be achieved by a coordinate translation.

An example of a Lie group is the group of $n \times n$ nonsingular matrices with real elements. This is an $n2$–dimensional group, in which the matrix elements, with 1 subtracted from the diagonal elements, can be regarded as the group parameters.

If two group elements a and b, with coordinates **p** and **q** respectively, are given a third element c their product ab is determined by the group multiplication property. The coordinates of c will be denoted by **p**. The numbers $r^1, r^2, \ldots, r^n$ depend on the coordinates of a and b and they are therefore functions of the variables **p** and **q**. There are n such functions, one for each coordinate of c, and each function depends on $2n$ variables, the coordinates of a and b. There seems to be no standard terminology for these functions; we will call them *product functions* and denote them by **f**. The coordinates of $c = ab$ are therefore given by **f**(**p**, **q**). For a given pair **p**, **q** there may be more than one function **f** since $c = ab$ may have coordinates in more than one parameter domain. The various possible functions are, because of B, connected by a differentiable transformation.

The group multiplication laws (2.1) and (2.2) impose certain conditions on the functions f. For example, the property $ae = ea = a$ requires that

$$\mathbf{f}(\mathbf{p}, \mathbf{0}) = \mathbf{f}(\mathbf{0}, \mathbf{p}) = \mathbf{p}. \tag{3.1}$$

The associative law (ab)c =a(bc) imposes a rather more complicated condition on the **f**. We again denote the coordinates of a, b, and c by **p**, **q** and **r** respectively. The group element ab has coordinates **f**(**p**, **q**) and $(ab)c$ therefore has coordinates **f**(**f**(**p**, **q**), **r**). In a similar way it can be found that $a(bc)$ has coordinates **f**(**r**, **f**(**p**, **q**)). The equality of the two group elements requires that

$$\mathbf{f}(\mathbf{f}(\mathbf{p}, \mathbf{q}), \mathbf{r}) = \mathbf{f}(\mathbf{p}, \mathbf{f}(\mathbf{q}, \mathbf{r})). \tag{3.2}$$

This relation will be called the associative law, since it is equivalent to (2.1).

The differentiability condition for the group operations is formulated in terms of the functions **f**(**p**, **q**) and is that they must have uniformly bounded derivatives of all orders with respect to the variables p^α and q^β.

The meaning of the closure property imposed on a Lie group is that if any sequence $\{\mathbf{p_n}\}$ of group coordinates in some parameter domain D converges to a point **p**, the group elements g_n corresponding to $\mathbf{p_n}$ converge to a group element g. This is actually the definition of the convergence of group elements. This condition is significant only if **p** is on the boundary, and hence outside of D. The meaning is then that there must be a second parameter domain D' in which, at least past a certain point in the sequence, the group elements g_n have coordinates $\mathbf{q_n}$, and the

$\mathbf{q_n}$ converge to $\mathbf{q}$, an interior point of D'. This assumption is a condition on both the group and the admissible parameterizations.

The nature of this assumption can perhaps be clarified by the following example. The group R of real numbers, with addition playing the role of group multiplication, is closed in the usual parametrization. A new parameter y =tan-1 x could be introduced for which -1f/2 ¡ Y¡ 1f/2. This parametrization is inadmissible, since there is no group element corresponding exactly to the limit elements at y = 1f/2. The group of rotations in the plane is closed although there is apparently no limit element at 1f fo r the parameter () above. The rotation by 1f corresponds, however, to an interior point in the ()2 coordinate system, thereby satisfying the condition given above. The property that the group is connected is that the coordinates (in at least o1).e parameter domain) of every group element can be connected to the coordinates of the identity, in the parameter domain in which they are zero, by a continuous curve. This curve can pass from one parameter domain to another, but in doing so must pass continuously through the region of overlap of the two domains. If any two such curves, joining the identity to an arbitrary group element, can be deformed continuously into one another, the group is simply connected; otherwise it is multiply connected. These properties do not, however, play an important role in the work that we are considering here.

3.2 Compact Groups

The theory of representations has particularly simple results for a certain class of groups, the compact groups. The property of compactness also simplifies the derivation of some general theorems. In this section we will define a compact group. A simple example of a compact (not group theoretical) set is a closed finite interval on the real line. Such an interval, [a, bJ, has two important properties. The first is that for any infinite set of points , a :s Xi :s b, there is (at least) one limit point also in [a, b]. If every neighborhood of a point x contains an infinite number of the points Xi, x is said to be a *limit element* of the set xi' The second property of a closed finite interval could be called the Heine-Borel property. This stipulates that from any set of open intervals Ia = (xa , Ya) that covers [a, b], it is possible to select a finite subset of intervals Ia1 ' Ia2 ' ??? , Ian that again cover [a, b]. The verb "cover" means in this case that each element of [a, b] is also in one of the covering open intervals. The classic Heine-Borel theorem is that the second property is a consequence of the first property. A Lie group will be defined to be compact if there is a parametrization in which the group is covered by coordinates in a finite number of bounded parameter domains. It is essential, of course, that the parametrization be such that the differentiability and closure properties described above are satisfied. A bounded parameter domain is one that can be enclosed in a rectangle in the coordinate space. Specifically, there are numbers ai, bi, i = 1, 2, ..., n such that It ai ¡ pi ¡ bi for the coordinates pi of points in the domain. Let be an infinite set of elements of a compact Lie group.)int 1 3 ain ,s any t, Y 10 Then one of the parameter domains must also contain the coordinates of an infinite number of the gn and these coordinates must have a limit element either in or on the boundary of the parameter domain. (The proof can be found in many books on advanced calculus.) The closure property of the group then assures that the limit element is a group element. The first property of closed finite intervals, that any infinite sequence of group elements has a limit element, is therefore satisfied by a compact Lie group. It is also possible to adapt the usual proof of the Heine-Borel theorem to show that a compact Lie group G has the second property of closed finite intervals in the following sense. Suppose a family of open regions Ua covers G by including the coordinates of each element of G from at least one parameter domain. It is then possible to select a finite number of the Ua that cover G in the same way. (The definition is complicated here by the possibility that a given group element may have coordinates in several parameter domains.)

It is obvious that the group of rotations in the plane is compact since it is covered by two finite intervals. On the other hand, the group R of real numbers cannot be compact since the sequence of the integers 1, 2, ... cannot have a limit element in any parametrization.

3.3 Locally Isomorphic Groups

The theory of representations has particularly simple results for a certain class of groups, the compact groups. The property of compactness also simplifies the derivation of some general theorems. In this section we will define a compact group.

It may happen that two groups which are-essentially different appear to be identical if one is permitted to view only a part of them. We will give a rather simple example of this situation. We consider again the group R of all real numbers with addition being the operation of group multiplication. This group has a normal subgroup, the group Z of all integers. The factor group R/ Z, whose elements are cosets of Z, consists of sets of numbers such that the difference of any two of the numbers in a set is an integer. Each element can, therefore, be characterized by a number p satisfying -1/2 ¡ p ::s 1/2. The product of two elements of R/Z corresponding to p and q correspondstop+qif-1/2¡P+q ::s1/2;ifP+qisoutsidethisinterval it can be reduced by 1 to bring it back into the interval. The groups R and R/Z are certainly different; for example, R/Z is compact (it is isomorphic to the group of rotations in the plane) whereas R is not. The two groups are identical, however, if one considers only elements corresponding to Ipl ¡ O.I. The groups R and R/Z are s aid to be loc ally isomorphic. This property can be defined as follows. Lie groups G and G' are locally isomorphic if there exist neighborhoods of the identity in each group and parametrizations such that the product functions are equal if their arguments are in these neighborhoods.

3.4 Properties of the Product Functions

The following discussion will contain the calculation of derivatives of the product functions at various points in the group. The customary notation for partial derivatives can be rather misleading, so that we will introduce and adhere to the following convention. The symbol Ofa/or!!will be used to represent the derivative of the product function fa with respect to the 11th coordinate of the first set of arguments of fa, regardless of the point at which this function is actually evaluated. Similarly, ofa /8 q Y denotes the derivative of fa with respect to the yth coordinate of the second argument. The summation convention of summing over repeated indices will also be used. We will now discuss a few properties of the derivatives of the product functions. The first of these are rather trivial and are obtained by differentiating (1) with respect to p f3:

$$a + b \tag{3.3}$$

These imply further that

$$b + c \tag{3.4}$$

If rO' =P (p, q), it is possible to show that the Jacobian

$$c + d$$

is nonzero. The parameters of the inverse of the element with the r ,r ,??.,r) of0' 12 =–(pq) parameters p will be denoted by i1 (p), i2(p), ??., in (P) so that

$$d + e \tag{3.5}$$

Consider then the derivative with respect to rJ3 of the identity fO'(f(p,q),1(q=pO', which is a reflection of the identity abb-1 = a. The derivative is

$$e + f \tag{3.6}$$

This shows that the matrix whose elements are of Y/op(3 has an inverse and hence that the Jacobian is nonzero. It is possible to show in a similar way that

$$f + g$$

These results are simply reflections of the fact that the transformations r(p) = f(P, q) with q fixed, and r(q) = f(P, q) with p fixed, are necessarily invertible. It is now possible to calculate the derivatives of the functions 1(P) that are determined bl the inversion operation. Consider the derivative with respect to p of the identity (5):

$$e = ma$$

This equation can be solved for oi Y/op(3 by using equation (6) which instructs us how to invert the matrix whose elements are of0'/ op(3 ? The desired result is

$$2 + 2 \tag{3.7}$$

It is apparent that higher derivatives of i could be calculated by differentiation of this result. In the next section we will have occasion to use functions vO!f3 defined by

$$2 - 2 \tag{3.8}$$

These functions satisfy an identity which can be derived by differentiating equation (2), the associative law, with respect to r" and then setting r =O. Using the chain rule and (1) gives

$$2 + 1$$

or, from (8),

$$a - c \tag{3.9}$$

This equation will play an important role in Chapter 5 when an invariant integral over the group will be defined.

3.5 One Parameter Subgroups

The investigation of the structure of a Lie group G is facilitated by the construction of the one-parameter subgroups of G. A one-parameter subgroup is a "curve" a(t) in the group with the property that

$$a(s) + a(t) = a(s + t)$$

The variable t is a real number that serves to label the group elements in the subgroup; it is the "one-parameter." Corresponding to an element a(t) . G, there is a point p(t) in a parameter domain. An additional requirement on the one-parameter subgroup is that the points p(t) constitute one or more continuous curves in each parameter domain through which it passes. A rather trivial one parameter subgroup is defined by a(t) = e. If a(t) is a one-parameter subgroup, a(O!t) is also one consisting of the same group elements but differing in the scale of the parameter. It is evident from (**??**) that $a(0) = e$ and that $[a(t)]^{-1} = a(-t)$. The

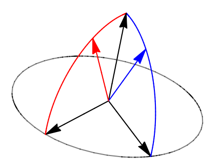

Figure 3.1. *A one-parameter subgroup showing a, the tangent at e, and $\dot{p}(t)$, the vector into which $v^{\alpha}_{\beta}(p(t))$ transforms a.*

one-parameter subgroup can be shown to be isomorphic to either R, the group of real numbers under addition, or to R/Z, the factor group discussed in Section 3 of this chapter, depending on whether or not there is a number t1 such that a(t1) =e. It has been implied that the one-parameter subgroups may pass through several parameter domains. It may also cross the boundary of a parameter domain and reenter it at another point. We shall first assume the existence of a one-parameter subgroup and obtain some of its properties. It will then be shown that there are one-parameter subgroups and that in a sufficiently small neigh- borhood of e, each.element is contained in a one-parameter subgroup. We restrict our attention now to the parameter domain containing the coordinates of e (which is the origin in the coordinate system of that domain). If it is assumed that the coordinates p(t) are differentiable functions, a first-order differential equation can be derived for the subgroup. The tangent vector to p(t) at t =0 will be denoted by a; that is

$$a = \dot{p}(0) \tag{3.10}$$

From equation (10) and the definition of the product functions f we can write

$$1 + 1 \tag{3.11}$$

at least for sand t close enough to 0 so that p(s), p(t), and p(s + t) remain in the original parameter domain. A differential equation for p(t) is obtained by differentiating (12) with respect to t,

$$pas$$

and setting t = O. Using (8) and (11), one obtains

$$pas \tag{3.12}$$

Equation (13) shows that the matrix whose elements are va,s (p) transforms the tangent vector at e of a one-parameter subgroup into the tangent vector at p, as shown in Fig. 3-1. It should be noted that, according to (3) and (8),

$$vab \tag{3.13}$$

This is necessary if (13) is to be consistent at e. It is also observed that matrix vj, (P) is nonsingular, since it is a special case of the matrix ofajoq,s. The initial condition pa (0) =0 can be added to the system of equations (13). A standard result, the Picard-Lindelof theorem [1], of the theory of ordinary differential equations guarantees that if the functions va (p) a,s have bounded derivatives with respect to their arguments (itepY in this case), then there is an interval r-e, e) I::¿uch that the equations (13) have a unique solution for s in this interval. The differentiability condition is satisfied in the present case since the product functions are at least twice differentiable in the parameter domain. We consider now a solution of (13), satisfying p(O) =0, for some particular choice of the a,s. It is to be expected that these solutions will also satisfy (12) and therefore describe at least part of a one-parameter subgroup. The components of the tangent at e to the subgroup will then be the a,s. This is the content of the following theorem.

Theorem 3.1. *A set of functions p(s) that satisfy* (3.12) *for s in the interval r-e, e) with initial condition p(O) = 0 also satisfy* (3.5) *for s, t, and s + t in the same interval. These functions can be used to generate a complete one-parameter subgroup.*

Proof. The method of proof is to show that each side of (3.5), considered as a function of t, satisfies the same first-order differential equation with the same initial conditions; the uniqueness theorem for the solution then guarantees that the two sides are the same. The left-hand side of (12) will be denoted by la(t); differentiating with respect to t and using (13) gives

$$1 \tag{3.14}$$

This equation is in fact identical to (13) but with the initial condition l(O) = p(s). The right-hand side of (12) will be denoted by r a (t). Differentiating r with respect to t and using successively (13), (9), and the definition of r gives

$$3-16 \tag{3.15}$$

This equation is identical to (15) so that it can be concluded that r(t) and l(t) satisfy the same differential equation. Furthermore, r(O) =f(p(s), 0) =p(s) so that rand l satisfy the same initial condition. The uniqueness theorem for the solution of the differential equations guarantees that r(t) and l(t) are the same in any interval in which the solutions exist, that is, provided s, t, and s + t are all in (- ,) since the solution of (15) or (16) is then obviously p(s + t). If a solution of (13) is known in the interval [-,], which can be assumed to be closed, it can be continued outside the interval by expressing an arbitrary parameter t in the form n + 0 where n is an integer and 0 :s 0 ¡ . The group element a(t) can then be defined to be [a()]na(o). It can be seen that the function a(t) defined in this way is a one-parameter subgroup since the various factors involved in the construction all commute. This subgroup is constructed in the group; a particular parameter domain may contain the coordinates of only a limited part of the subgroup. Henceforth, when we refer to the coordinates p(t) of a one-parameter subgroup, we mean the arc connected to the origin in the domain that contains the coordinates of the identity. It is also convenient to call p(t) one-parameter subgroup, even though it may really only be part of one. □

3.6 Canonical Coordinates

It has been seen that a solution to (3.12) exists for every tangent vector a so that a one-parameter subgroup passes through e in every direction. It will now be shown that there is a neighborhood V of e such that a one-parameter subgroup passes through each point in V. It is first noted that if p(t) is a one-parameter subgroup with tangent vector a, p(ct) is a one-parameter subgroup with tangent vector ca. We will now show the dependence of p(t) on a explicitly by writing p(t) = z(a, t) where a = p(O). Then the above observation shows that z(a, ct) = z(ca, t). Consider now the function w(a) = z(a, 1). It can be from the above identity that zO!(a, t) = wO!(ta). If this identity is differentiated -with respect to t we find

$$partial$$

from the differential equation (13). The theory of ordinary differential equations guarantees that the derivatives 8 wCJaaf3 exist. Putting t =0 now shows that

$$partial$$

and, since the numbers af3 can be chosen arbitrarily, that

$$partial$$

It can be concluded that the Jacobian of the transformation w(a) is 1 at a = O. The implicit function theorem can now be used to prove that there are open regions V in the parameter space and U in the space of tangent vectors such that the equation p =w(a) can be solved to give a as a single-valued function of p. In other words, for every p V, there is a unique a U such that p =w(a). It is therefore possible to use the components of the tangent vectors in U as a new set of group parameters characterizing the group element with old coordinates p =w(a), provided p V. The new coordinates of a group element are the components of the tangent at e to the one-parameter subgroup that passes through the point at t =1. This new coordinate system is said to be *canonical.*

The one-parameter subgroups have the very simple parametric equations a(t) =at as long as at is in U. This is proved as follows. Let p(t) be the old coordinates of a one-parameter subgroup with tangent vector a U; then p(l) has new coordinates a. Consider the curve ps in the old parameters defined by ps(t) =p(st). The tangent at 0 to ps is sa and ps (1) has new coordinates sa (provided sa U). On the other hand ps (1) = p(s), showing that the new coordinates of p(s) are simply sa, provided p(s) V (or sa U).

It is also possible to prove that any continuous one-parameter subgroup (the arc going through the unit element), not assumed differentiable, must have a parametric equation of the form p(t) =at for some a[2]. Therefore, the coordinates must actually be differentiable functions of t in the open region V and must satisfy equation (13) ? The proof of this fact will be omitted.

This has the following consequence that will be useful in the next chapter. Suppose that f is a homomorphic mapping of a Lie group G onto another Lie group G'. If a(t) is a one-parameter subgroup in G, f(a(t is a one-parameter subgroup of G'. Iffurther, the coordinates of G and G' are canonical, the coordinates of a(t) and f(a(t are expressible as at and a't respectively. It will be assumed that a and a' are in the neighborhoods V and V' in which canonical coordinates for G and.G' are defined; this can always be achieved by a change of scale in the parameter t. If f is regarded as a mapping from the coordinates of G to the coordinates of G' , we can write

$$fat3-17 \tag{3.16}$$

where a,a is independent of t but depends on a. It can be proved that the fa are linear functions of their arguments. It is first necessary to prove that they are differentiable at the origin. To show this we put Jl = oJJ.v so that (17) becomes

$$1$$

Differentiation with respect to t shows that afajOav (0) exists and is a,a(O, 0, ...,1, ...,0), the 1 being the vth argument of a,a. The derivative of (3.16) with respect to t, evaluated at t = 0, is

$$dfa$$

On the other hand, if we put t =1 in (17), we find that

$$faa$$

Equation (18) shows that f is a linear functions of a.

The transformation to canonical coordinates from an arbitrary parametrization is differentiable and invertible in the neighborhood of the origin. From this it can be shown, in view of the differentiability of (3.6), that a homomorphic mapping of one Lie group onto another must be a differentiable mapping of the coordinates in the neighborhood of the origin. In particular, two parametrizations in a Lie group must be connected by a differentiable transformation. In this case the transformation must also be invertible, since in each parametrization there are neighborhoods of the origin in which the points can be placed in one-to-one correspondence with the group elements, and therefore with each other.

3.7 One Parameter Subgroups of Matrix Groups

To provide an illustration of the concept of one-parameter subgroup, these subgroups of the group L(n) of n x n nonsingular matrices will be calculated. A parametrization of grq"!lP is obtained by expressing each matrix element in the form 0 1) + pI) so that the coordinates of the identity are zero. Rather than attempt to use a single index, it is much more convenient to use a pair of indices, namely, (ij), i = 1, 2, ... , n, j = 1, 2, ... , n to label the parameters. The product functions for the matrix group are then given by

$$3-19 \tag{3.17}$$

From this we obtain

$$3-20 \tag{3.18}$$

The differential equation (**??**) is therefore

$$3-21$$

where the amn are the matrix elements of the "tangent" matrix A. c This equation can be conveniently written in matrix form as

$$\dot{P} = A + PA \tag{3.19}$$

It can be verified that this equation has the solution

$$3-23$$

where

$$3-24$$

is a convergent (for all t) power series solution of (**??**). (The largest matrix element of Ar.is smaller in absolute value than (nAm)r where Am is the absolute value of the largest matrix element of A.) We can conclude that the one-parameter subgroups of L(n) can be written in the form eAt where A is an arbitrary n x n matrix.

Chapter 4

Lie algebras

The properties of a Lie group that pertain to elements in the neighborhood of the identity can be investigated and characterized by considering another mathematical structure associated with the group, the Lie algebra of the group. The purpose of this chapter is to define Lie algebras, and to discuss, rather superficially, the relation between a Lie group and the corresponding Lie algebra. The Lie algebra of an arbitrary Lie group will be defined in Section 1, and in the second section this definition will be related to the Lie algebra of a matrix group, which can be defined, somewhat differently, as a matrix algebra. Further properties of Lie algebras are developed in Sections 4, 5, and 6. In Section 7 the Lie algebras of some of the important matrix groups are calculated, and in Section 8 all the Lie algebras of dimension 1, 2, and 3 are catalogued. Some of the theorems which are more difficult to prove will be stated without complete proof in this chapter. These theorems concern the relation between Lie groups and Lie algebras; their proofs can be found in Pontrjagin's classic work [1]. The term Lie algebra was introduced by H. Weyl to replace the earlier somewhat misleading term, infinitesimal group, used by Lie. 4-1 ABSTRACT DEFINITION OF A LIE ALGEBRA The elements of a Lie algebra are real, n-dimensional vectors a, and a Lie algebra A is an n-dimensional vector space. A Lie algebra is constructed from a Lie group G by defining the vectors a to be the tangents at the identity to (at least twice) differentiable curves in the group These curves may, but need not, be one-parameter subgroups. In any event, for any a e A there is, according to Theorem 3-1, a one-parameter subgroup p(s) such that in . P(O) = a. The linear properties of A are related to the group properties of G. If, for example, a and b are tangent to curves p(s) and q(t) respectively in the group, a + b is tangent to the curve r(s) = p(s) q(s). The parameters of r(s) are given by rqs) =fD'(p(s), q(s; therefore a 8fa .3 8fa '3 a a r (0) =8p3 (O,O)p (0) + 8q3 (O,O)q (0) =a + b because of (3.3) and (1). If a is tangent to the curve p(s), Q'a is gent top(as). In particular, if p(s) is a one-parameter subgroup, and a is an integer, Q'a is tangent to [p(s)]a. In a very loose sense, A can be regarded as the space of "logarithms" of group elements. In an algebra, not only linear combinations, but also products of elements are defined. In the case of a Lie algebra, the product of two elements is called the commutator and is constructed from the group commutator., Suppose a andb are tangent to twice differentiable curves p(s) and q respectively. It is convenient, but not necessary, to assume that p(s) and q(t) are one-parameter subgroups. We form the function a(s,t) =P(s)q(t)p(s)-iq(t)-i, (4.2) where p(s), q(t) represent the group elements with these parameters . Equation (2) determines, in general, a surface in the group. The parameters r of a(s, t) are given by (4.3a) LIE where Th whOSE The (const Tl to sir wew: choic Tl fined Al tor t defin T as a THE met] and E q u r a (s, t) = ra(f(p (8), q(t, f(p (-8), q (-t) if p(s) and q(t) are one-parameter subgroups. We consider expanding r a(s, t) to second-order in a Taylor series in sand t: a a a a .a2 at at2 r (8,t) =ao + 8 + a2 t + a11 s + 8 + a22 + ???? (4.3b) This is possible since p(s), q(t) , and also the product functions are at least twice differentiable. It can be seen from (2) that a(O, t) = e for all t since p(O) = p(O)-1 =e and hence that ra(O,t) =0. Putting s = in (3b) shows that aoa = a2a = a22a = 0. Similarly, it can be shown by putting t = 0 that a1a =a11a =0, so that r a (8 , t) = + ? ? ? ? The coefficients a12Q' can be calculated from equations (3): 82ra [82fa 82fa J. . = 8s8t (0,0) = 8p38q 'Y (0,0) - 8q38p'Y (0, O)J p3 (O)q'Y (0) (4 . 3 c) =ca b 'Y where a c (3y = apf3aqY (0,0) -

apYaq3 (0,0). 29 (4.4b) .4a) The commutator [a, b] of the elements a and b of A is the vector whose components are defined by ba a a 3Y [a,]= =C3yab? (4.5a) The coefficients ca (3y defined in (4b) are known as the structure constants of the Lie algebra. The choice of p(s) and q(t) as one-parameter subgroups was made to simplify the discussion and is not essential. It can be shown, but we will not do this, that the definition of [a, b] is independent of the choice of curves to which a and b are tangent. The unit vectors in A, that is, the vectors ea with components defined by (ea)fJ = 0afJ, can be seen to satisfy (4.5b) Another definition of the commutator is that it is the tangent vector to the curve a(rs, -IS), as defined by (2). The advantage of this definition is that it is independent of the parametrization in the group. Two simple consequences of equations (4b) and (5a) can be stated as a theorem. THEOREM 4-1. The commutation operation satisfied the antisymmetry property and the linearity properties Equations (6 imply that [a, a] = 0 and that cafJfJ = o. (4.6) (4.7a) (4.7b) The structure constants depend on the parametrization chosen for the group G. Under a change in parametrization of G, the vectors ea of equation (5b), which are unit vectors along the coordinate axes , must be expressed in terms of ffJ, the unit vectors along the new coordinate axes. Explicitly, ea = sa(3f[, f(3 = t(3a ea , where sa(3 = op(3lopa (0) and t(3a = opa lop(3(o). I cPaT are the structure constants in the new representation, we can write (3y -t[3YO! [fa,fTl=ta tT [e(3,eyl- a tT c (3yeO! = ta (3t T YsO!Pca(3yfp =cP aTfp' This shows that in the new parametrization the structure constants are given by (4.8) If two Lie algebras are of the same dimension and their structure constants are the same or can be connected by a transformation of the form (8), where the t af3 are the matrix elements of the inverse of the matrix whose elements are SO! P, the Lie algebras are said to be isomorphic. All isomorphic Lie algebras are essentially the same; the properties of one of them imply all the algebraic properties of the others. If the commutator of any pair of elements in a Lie algebra is zero, the algebra is said to be Abelian. 4-2 LIE ALGEBRA OF A MATRIX GROUP If a Lie group is a matrix group, it is possible to define its Lie algebra not as tangents in the parameter space, but as a set of matrices. The result is a Lie algebra isomorphic to that of the definition of the previous section, that is, it is a vector space of tL", same dimension and has a commutation operation defined with the same structure constants. The commutation operation is in this case, moreover, the operation of forming matrix commutators, that is, [A"B] =AB - EA. (4.9) We consider a Lie group G of matrices A(p) with the group operation being that of matrix multiplication. It will be assumed that the elements of A are differentiable functions of the parameters p. Suppose that a =p(O) is an element of the Lie algebra A of G; a is the tangent at 0 to a curve p(t) in the parameter space. Corresponding to the curve p(t) there is a matrix function A(p(t. We can calculate A(P(O)) = oAO! (0) pO! (0) = IO!aO! (4.10) op where It will be assumed, although it can be proved on the basis of the previous axioms, that the IO! are linearly independent. This will be evident in all the examples that are considered. There is then a nonzero matrix 100aO! corresponding to each a [, A. The set A' of such matrices is an n-dimensional vector space, which will be shown also to be a Lie algebra. The matrices IO! are called the generators of AI. It will now be shown that Ipl 'Y - lylj3 is also in Al and is, in fact, cO!,B'YIO! where the cO!,B'Y are the structure constants of A. This is shown by applying equation (2) directly to G, with p(s) and q(t) assumed to be one-parameter subgroups:

1.9) ra- le Ip- ; . The desired result can be obtained by evaluating the second derivative, with respect to sand t, of this identity, at s =t =O. It is simpler, however, to expand each side to second order in sand t and equate the coefficients of st on each side. Since we have seen that terms in s2 and t2 do not appear in equation (3b), it is sufficient to expand A(P(s and A(q(t to first order in s and t, that is, A(P(s =I+ (O)a,Bs+???=I+l,Ba,Bs+.... Similarly,A(q(t =1+l'Yb'Yt+.... To first order in sand t, A(P(-s =1-I paPs, and A(q(-t =1- 1'Yb'Yt. In the desired approximation we can also write rO!(s, t) =cO!p'Ya fib 'Y st and therefore A(r(s, t =1+ 1O!cO!fi'Ya fib'Y st. Equa- tion (12) is therefore, in second order,

$$a$$

Equating the coefficient of st on each side shows that

$$b$$

The numbers a fl and b I' can be chosen arbitrarily, and therefore

$$b \tag{4.1}$$

Comparison of this result with (5b) shows that the basis vectors 10' in A' satisfy the same relations, under the product defined by (9), as the basis vectors eO' in A. Furthermore, the product in (9) is linear in the factors A and B so that if a 0' 10' and bfllfl correspond to a and b respectively, (aala)(bfl1fl) - (bflIfl)(aO'Ia) corresponds to [a, b]. Explicitly,

$$c$$

We can conclude:

Theorem 4.1. *The Lie algebra AI spanned by matrices Λ' defined by equation* (**??**) *with multiplication defined by* (**??**)*, is isomorphic to the Lie algebra Λ defined by equations* (**??**) *and (5a).*

It should be remarked that equations (**??**) and (**??**) depend on the possibility of differentiating the matrices with respect to their para- meters; this in turn is based on the possibility of taking linear com- binations (which are not defined in an arbitrary Lie group) of ma- trices. The resulting matrices, the 10', are not, in general, group elements.

4.1 Example

Before discussing further properties of Lie algebras, the rather abstract considerations of the last two sections will be illustrated by an example. We choose for this the affine group in one dimension, conSisting of matrices of the form =abc 0'BIY=[a,] II'. A(pi,p2) = (epi p2) o1' The law of group multiplication is given by f1(pt,p2;qi,q2) =pi + qt

The elements of the Lie algebra in the abstract sense are two-dimensional vectors a = (a1, a2). It can be verified directly that the subgroup associated with this vector consists of the matrices

$$A(s) = \begin{pmatrix} e^{a^1 s} & \frac{a^2}{a^1}\left(e^{a^1 s} - 1\right) \\ 0 & 1 \end{pmatrix} \tag{4.2}$$

The parameters of a(s, t) can be obtained by comparison with (**??**):

$$p^1 = 0 \tag{4.3a}$$

$$p^2 = 0 \tag{4.3b}$$

The element a(s, t) can be obtained by a bit of computation. 2 2'

The last line in (17b) give p2 correctly up to second-order terms in sand t. By equations (3c) and (5a) we conclude that (4.18) The one-parameter subgroup associated with this element of the Lie algebra is obtained from (16) by substituting 0 for a1 and a1b2 - a2b1fora2. It is, taking the limit , 1 (a1b2 -la2b1)S) A(s) = 0 The commutator could also have been obtained by calculating the structure constants from equation (4b). For this purpose only and C212 are needed; the remainder are either zero or can be obtained from equation (6). From equation (15) it is seen that and hence that 0 =-c121' Also, so that C212=-C221=1-0=1. From(5a) it is found therefore that [a,b1i =0 in agreement with the previous calculation. We have used the fact also that =C122 =C211 =C222 =O. The "more concrete" form of the Lie algebra makes use of the matrices Ia = 8A / 8pa(O). These are, from (14), I , C :) I , The general element of the algebra is a i ai It + a212 = (0 a 2) These matrices form a 2-dimensional vector space and at [ai I1 + a2I2,b1It + b2I2] =(0 ,a'b' a'b'), which corresponds to the [a., b1 of (18). Evidently the construction of the Lie algebra is most straightforward from the matrix definition. However, the definition in terms of tangent vectors is more natural for abstract groups and also permits a simpler discussion of some of the relations between Lie algebras and Lie groups. These relations will be discussed in Section 6 of this chapter. 4-4 THE JACOBI IDENTITY It has been seen that the commutation operation in a Lie algebra satisfies the antisymmetry and linearity properties of equations (6) and (7). These are not in themselves sufficient to define a Lie algebra; a third property, the Jacobi identity, must also be satisfied. The identity is: THEOREM 4-3 [a, [b,c]] + [C, [a,b]] + [b, [C,a]] = O. 0 The identity is a reflection of the associative law of group multiplication, although this fact is not obvious. If the Lie group is a matrix group, the a, b, c are matrices and the commutators in (19) can be evaluated from equation (9). In this case equation (19) is a consequence of the associative law of matrix multiplication as one finds by writing out the 3 x 2x 2 = 12 terms explicitly. Inasmuch as there is a matrix group locally isomorphic to any Lie group (though we have not proved this and will prove it in Chapter 6 only for groups that have a center conSisting only of e), equation (19) is a consequence of the simple proof for matrix Lie groups. Equation (19) can, however, be proved directly from the associative law (3.2) of group multiplication. The method of proof is to differentiate (3.2) three times, with respect to pJ3,qY, and r 0, and evaluate the result atp =q =r = o. The resulting identity can be rearranged to demonstrate (19). The calculation is, however, tedious and not instructive, and will not be given here. It should be noted that although the Jacobi identity is closely connected with the associative law of the group, the associative law need not be valid for the multiplication law of the algebra. Equation (19) provides a restriction on the structure constants. If a, b, c are basis vectors e3, ey, eo and equation (5b) is taken into account, the result is

$$e^{\alpha} \tag{4.4}$$

This is the O'th component of the identity. In (20) the indices 0', 3, y, 0 are free and J.1 is a summation index; the three terms are generated by cyclic permutations of (3, y, and o. If any two of these are equal, (20) reduces to an identity as a result of the antisymmetry relation (6). Hence in the previous example of a 2-dimensional Lie algebra, the Jacobi relations are automatically satisfied. Equation (20) is actually the result that is derived from the group associative law (3.2), rather than (19). An n-dimensional linear vector space with a product defined that satisfies the linearity, antisymmetry and Jacobi identities is also said to be a Lie algebra, whether or not it can be constructed from a Lie group. It was shown by Lie that for every such Lie algebra, there is a local-Lie group that has that Lie algebra. A local Lie group is essentially a set of functions f defined in some neighborhood of the origin that satisfy equation (3.1) and (3.2) and therefore define, for points p and q in the neighborhood of the origin, a product. The local Lie group may not be a Lie group, however, since f(P, q) may be outside the domain of the product functions, so that multiplication by f(p, q) is not defined. The proof of this result is quite involved and will not be given here. 19) The identity is a reflection of the associative law of group multiplication, although this fact is not obvious. If the Lie group is a matrix group, the a, b, c are matrices and the commutators in (19) can be evaluated from equation (9). In this case equation (19) is a consequence of the associative law of matrix multiplication as one finds by writing out the 3 x 2x 2 = 12 terms explicitly. Inasmuch as there is a matrix group locally isomorphic to any Lie group (though we have not proved

this and will prove it in Chapter 6 only for groups that have a center conSisting only of e), equation (19) is a consequence of the simple proof for matrix Lie groups. Equation (19) can, however, be proved directly from the associative law (3.2) of group multiplication. The method of proof is to differentiate (3.2) three times, with respect to pJ3,qY, and r 0, and evaluate the result atp =q =r = o. The resulting identity can be rearranged to demonstrate (19). The calculation is, however, tedious and not instructive, and will not be given here. It should be noted that although the Jacobi identity is closely connected with the associative law of the group, the associative law need not be valid for the multiplication law of the algebra. Equation (19) provides a restriction on the structure constants. If a, b, c are basis vectors e3, ey, eo and equation (5b) is taken into account, the result is 0' J.1 0' CJ.1 0' J.1 -0 c 3J.1c yo+c 0J.1 3y+c yJ.1c 03 ? (4.20) It is, furthermore, true that groups that have the same Lie algebra are locally isomorphic. It is from this fact that the Lie algebra obtains its importance; it determines all the properties of the elements of a Lie group in the neighborhood of the identity, that is, all the local properties. This is a consequence of the proof of Lie's theorem, in which the functions f are constructed in canonical coordinates as solutions of partial differential equations, and the uniqueness of the solutions of these equations can be demonstrated. 4-5 SUBALGEBRAS AND FACTOR ALGEBRAS It is possible to define, in analogy to group theory, ,a subalgebra of a Lie algebra. An m-dimensionallinear subspace M of a Lie algebra A is a subalgebra if, for each pair a and b in M, [a, bl is also in M. It is possible to formulate the existence of a subalgebra in terms of the structure constants. The basis vectors e;J- could, if M is a subalgebra, be chosen so that e10 e2, ... , em are in M. The conditionthatrei.ejl , MifeiandejareinMcanbeseentobe (4.21) It is apparent that M is a subalgebra if there is a transformation of the form (8) such that the transformed structure constants have the property (21). A subalgebra M is said to be an ideal if, for each a , Aand x , M, [a, xl , M. It will be seen that the ideal is the analogue of the normal subgroup of group theory and could perhaps be more suitably called a normal subalgebra. In the coordinate system in which the first m basis vectors are in M, the structure constants satisfy cO'B =0, 0' ¿ m, f3 :::; m (4.22) ,')I for all ')I. Again we remark that A contains an ideal if the structure constants can be transformed to have property (22). If a Lie algebra A contains an ideal M, it is possible to define a factor algebra AIM as follows. We consider the set a + M of all elementsoftheforma+m,m , M. Thissetisinfactagroup theoretic coset of M where M is a normal subgroup of A under the group operation of vector addition. Two elements a and a' are in the samecosetifandonlyifa- a', M.Thefamilyofallcosetsa+M is defined to be the factor algebra AIM. It is necessary to define the linear and commutation operations in A1M. These are given by c (a + M) = ca + M, (4.23

$$(a + M) \tag{4.5}$$

$$[a + M]$$

In order that these definitions be consistent, it is necessary to verify that they are independent of the choice of a and b from their respective cosets. We will do this for (25) and leave to the reader the corresponding arguments for (23) and (24). If a' and b' are arbitrary elements from a + M and b + M, we can write a' = a + m1 and b' =b+m2 where m1 and m2 are in M. In terms of a' and b',(25) becomes

6

since [a, m2], [mb b], and [mb m2] are all inM and m + M = M if m M. It is here that the property that M is an ideal is used. The zero element of AIM is the

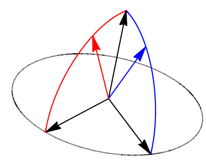

Figure 4.1. *A 2-dimensional Lie algebra with ideal M. The elements of AIM are lines parallel to M. In this case AIM is Abelian, since it is one-dimensional.*

coset M. It is easy to verily from (25) that antisymmetry and Jacobi conditions are satisfied in AIM. In Fig. 4-1 we show A as a plane and M as a line through the origin; the factor algebra AIM consists of all lines parallel to M.

A Lie algebra A is said to be homomorphic to a Lie algebra N if there is a linear function f mapping A onto N with the property that for any pair of elements a and b in A

$$f(a, b) \tag{4.6}$$

We note for completeness that a linear function is one with the property that

$$f(a, b) \tag{4.7}$$

and that this implies $f(O) = 0$.

If a Lie algebra A is homomorphic to a Lie algebra N, the set K zo of elements of A that map into the zero element of N is called the kernel of the homomorphism. That is, a , K if and only if f(a) [eO' ideal since if x , K, a , A, f([x, aD = [f(x),f(a)] = [0, f(a)] =0, and hence [x, a] , K. To complete the analogy to group theory, we remark that the factor algebra A/K is isomorphic to N by the correspondence if a ele: iso: adj' tati 4-E alg ob' im =C G' SUI eXI AT to wh aCe tia gh a +K-f(a). (4.28) If the kernel of f is 0, f has an inverse and is, therefore, an isomor- phism.

The set Z of all elements z of a Lie algebra with the property that [z,a] =for all a , A is called the center of A. It can be readily established that Z is a linear subspace of A and that Z is an ideal, since [z,a]= for all a, A and , Z.

The set of all linear combinations of elements that can be ex- pressed as commutators of elements in fis an ideal which is called the derived subalgebra of A and is denoted by A'; A' is an ideal since for any x , (I. I and a , A, [a,x] , AT by the definition. The derived subalgebra has the property that AIA' is Abelian. The proof of this fact is analogous to the proof of the similar result in the theory of groups and will be omitted.

If the center of a Lie algebra 1 consists only of 0, it is possible to obtain a Lie algebra of matrices isomorphic to /Ṡuppose ebe2, ??? , en are basis vectors in A in satisfying [e3 ,ey] = cO'3y eO'. We consider n x n matrices 11,12, ??? , In defined by

$$i = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \ j = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \ k = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.8}$$

Chapter 5

Invariant integration

Chapter 6

Group representations

Chapter 7

Completeness Theorems for Group Representations

Chapter 8

The Groups U(1) and SU(2)

Chapter 9

Rotations in space

The group $O(3)$ of rotations in a three-dimensional space is of great importance in mathematical physics. In this chapter, this group will be studied in general and the irreducible representations will be discussed. This leads to a development of certain properties of some of the special functions of mathematical physics; these are the Jacobi polynomials with their special cases of Legendre polynomials and associated Legendre functions.

The representations of the rotation group can be obtained easily by establishing a homomorphic mapping of $SU(2)$ onto $O(3)$; it can then be shown that certain representations of $SU(2)$ are also representations of $O(3)$.

The representations of $O(3)$ can also be constructed in invariant subspaces of the space of functions defined in three-dimensional space. It will be shown that the suitable invariant subspaces are composed of harmonic, homogeneous polynomials in three variables. In this way the connection between the group representations and Laplace's equation can be established. We will also consider and solve the partial differential equations that the representations are required to satisfy because of the group structure.

9.1 General properties of the rotation group

In this discussion we will denote by $\mathbf{R}$ any element of $O(3)$. The matrix $\mathbf{R}$ is a real 3×3 matrix that satisfies

$$\mathbf{R}^T\mathbf{R} = \mathbf{I}. \tag{9.1}$$

A consequence of (9.1) is that $\det \mathbf{R} = \pm 1$. The matrices with determinant -1 describe a rotation with a reflection; for the most part we will be concerned with the proper rotations, the subgroup of matrices of unit determinant. This group is denoted by $O(3)+$.

An arbitrary group element $\mathbf{R}$ has in general three eigenvalues which, since $\mathbf{R}$ is orthogonal, are of unit modulus. If $\mathbf{R}$ is a proper rotation the three eigenvalues of $\mathbf{R}$ must satisfy $\lambda_1\lambda_2\lambda_3 = 1$. Since the secular equation $|\mathbf{R} - \lambda\mathbf{I}| = 0$ is of degree three and has real coefficients, $\mathbf{R}$ has one real eigenvalue, which must be ± 1, and a pair of complex eigenvalues λ and λ^*. If the rotation is proper, the real eigenvalue must be $+1$. For a particular $\mathbf{R}$ it may happen that $\lambda = \lambda^* = +1$ or $\lambda = \lambda^* = -1$. The first case is clearly $\mathbf{R} = \mathbf{I}$; it will be seen that the second case corresponds to a rotation about some axis by π.

If $\mathbf{R}$ is a proper rotation there is a vector v with the property $\mathbf{R}v = v$ corresponding to the eigenvalue 1. A line through the origin in the direction of v is invariant under the rotation $\mathbf{R}$ and can be interpreted as the axis of rotation. If $\lambda = e^{-i\phi}$, ϕ real, is another eigenvalue there is associated with it a complex vector $u + iw$ satisfying

$$\mathbf{R}\,(u + iw) = e^{-i\phi}\,(u + iw) \tag{9.2}$$

The vectors u and w are each perpendicular to v since v and $u+iw$ are eigenvectors corresponding to different eigenvalues implying $(u + iw) \cdot v = 0$. It is possible to show also that $u \cdot w = 0$. If the complex conjugate of (9.2) is taken, it is seen that, since $\mathbf{R}$ is real, $u-iw$ is also an eigenvector of $\mathbf{R}$ corresponding to the eigenvalue $e^{i\phi}$. The orthogonality of these two vectors (in the usual complex vector inner product) implies that

$$(u + iw) \cdot (u + iw) = |u|^2 - |w|^2 + 2iuw = 0$$

This result shows that $|u| = |w|$ and $u{\cdot}w = 0$. The vectors u and w can be visualized as mutually perpendicular vectors lying in a plane perpendicular to v.

Taking real and imaginary parts of (9.2) shows that

$$\mathbf{R}u = \quad \cos\phi\, u + \sin\phi\, w \tag{9.3a}$$

$$\mathbf{R}w = -\sin\phi\, u + \cos\phi\, w \tag{9.3b}$$

This result indicates that vectors in the plane perpendicular to v, which can be expressed as a linear combination of u and w, are rotated within the plane by an angle ϕ. An arbitrary vector $\mathbf{R}$ can be expressed in the form $\alpha v + \beta p$ where p is in the plane perpendicular to v. We can write $\mathbf{R}r = \alpha v + \beta \mathbf{R}p$, indicating that $\mathbf{R}$ has been rotated about v by an angle ϕ In Fig. 9-1 we show the vector v, the plane spanned by u and w, and an arbitrary vector r together with $\mathbf{R}r$.

If the vectors u, w, v, are assumed to be normalized to unit length the matrix $\mathbf{Q}$ whose columns are u, w, v, in that order, is orthogonal. It can be verified that, since $\mathbf{Q}^T = \mathbf{Q}^{-1}$,

$$\mathbf{Q}^{-1}\mathbf{R}\mathbf{Q} = \mathbf{R}' = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9.4}$$

The matrix $\mathbf{R}'$ defines a right-hand rotation about the z axis by ϕ. Since $\mathbf{Q} \in O(3)$, $\mathbf{R}$ and $\mathbf{R}'$ are in the same class and we can conclude that a rotation about any axis by an angle ϕ is in the same class as a rotation about the z axis by ϕ. This implies that if two rotations are by the same angle they are in the same class. On the other hand, rotations by different angles are in different classes since their diagonal forms are different. It can be noted that since the trace is invariant under the transformation (9.4) the angle of rotation can be determined from

$$\mathrm{tr}\ \mathbf{R} = 1 + 2\cos\phi. \tag{9.5}$$

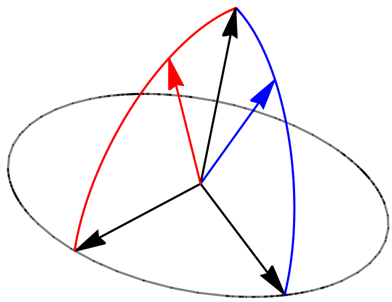

Figure 9.1. *The effect on an arbitrary vector r of a rotation about the v axis by an angle ϕ.*

The group can be parametrized by specifying the polar coordinates of v, the axis of rotation, and ϕ, the angle of rotation, where $0 \leq \phi \leq \phi$. A more common parametrization, however, is by the **Euler angles** which will now be defined. A rotation is uniquely specified by the final position of three unit vectors, i, j, k, which were originally parallel to the x, y, and z axes respectively. The final components of k can be written $(\sin\beta\sin\alpha, -\sin\beta\cos\alpha, \cos\beta)$ where β is the colatitude of k and $\alpha = \phi + \pi/2$, where ϕ is the azimuthal angle of k. In fig. (9.2) we show the final position of k and the angles β and α.

Figure 9.2. *The vectors e, f, k show the position of i, j, k following the rotation $Z(\alpha)X(\beta)$. Note that e is in the $x-y$ plane perpendicular to k and that $f = k \times e$. The angles β and $\alpha - \pi/2$ are the spherical polar coordinates of k.*

The vectors i and j lie in the plane perpendicular to k; the rotation can be completely determined by specifying the orientation of i and j in this plane. If two unit vectors in this plane are known, i and j can be expressed as a linear combination of them. One such vector is the vector e lying in the $x-y$ plane with components $(\cos\alpha, \sin\alpha, 0)$ and another, perpendicular to both k and e is $f = kxe$ with components $(-\sin\alpha, \cos\alpha\cos\alpha, \sin\beta)$. The vectors i and j can be expressed uniquely in the form

$$\begin{aligned} i &= \quad e\cos\gamma + f\sin\gamma, \\ j &= -e\sin\gamma + f\cos\gamma. \end{aligned}$$

The angles α, β, γ defined in this way are the Euler angles; it is observed that these angles characterize the rotation completely. The domain of the angles is $0 \leq \alpha < 2\pi$. $0 \leq \beta < \pi$, $0 < \gamma < 2\pi$. The vectors e and f are also shown in fig. (9.2) and in fig. (9.2) the vectors i and j are shown in the final position.

Figure 9.3. *The final position of* i*,* j*,* k *following the rotation* $Z(\alpha)\,X(\beta)\,Z(\gamma)$ *showing* i *and* j *rotated in the* $e-f$ *plane relative to* e *and* f *and* γ*.*

The vectors i and j have components

$$(\cos\alpha\cos\gamma-\sin\alpha\cos\beta\sin\gamma,\cos\alpha\cos\gamma-\sin\alpha\cos\beta\sin\gamma,\sin\beta\cos\gamma)$$

and

$$(-\cos\alpha\sin\gamma,-\sin\alpha\cos\beta\cos\gamma,-\sin\alpha\sin\gamma+\cos\alpha\cos\beta\cos\gamma,\sin\beta\cos\gamma)$$

respectively. The matrix that transforms the initial components of i, j, and k to the final components has for its columns the components of i, j, and k in their final position. The matrix $\mathbf{R}$ with Euler angles α, β, and γ is, therefore,

$$\mathbf{R}(\alpha,\beta,\gamma)=$$
$$\begin{pmatrix}\cos\alpha\cos\gamma-\sin\alpha\cos\beta\sin\gamma & -\cos\alpha\sin\gamma-\sin\alpha\cos\beta\sin\gamma & \sin\beta\sin\alpha\\ \sin\alpha\cos\gamma+\cos\alpha\cos\beta\sin\gamma & -\sin\alpha\sin\gamma+\cos\alpha\cos\beta\cos\gamma & -\sin\beta\sin\alpha\\ \sin\beta\sin\gamma & \sin\beta\cos\gamma & \cos\beta\end{pmatrix} \tag{9.6}$$

It can be verified by a direct calculation that the matrix $\mathbf{R}(\alpha,\beta,\gamma)$ is equal to the matrix $Z(\alpha)\,X(\beta)\,Z(\gamma)$ where

$$Z(\alpha)=\begin{pmatrix}\cos\alpha & -\sin\alpha & 0\\ \sin\alpha & \cos\alpha & 0\\ 0 & 0 & 1\end{pmatrix},\quad X(\beta)=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\beta & -\sin\beta\\ 0 & \sin\beta & \cos\beta\end{pmatrix} \tag{9.7}$$

The matrix $Z(\alpha)$ is a rotation about the z axis by α and the matrix $X(\beta)$ is a rotation about the x axis by β. The reason for this result can be explained as follows. Consider first the rotation $Z(\alpha)X(\beta)$. The rotation $X(\beta)$ rotates k within the $y-z$ plane to have colatitude β (and azimuthal angle -$\pi/2$) and leaves i unchanged. The rotation $Z(\alpha)$ then rotates k about the z axis leaving the colatitude unchanged but increasing the azimuthal angle to $\alpha-\pi/2$; $Z(\alpha)X(\beta)$ therefore, rotates k to its final position. The rotation $Z(\alpha)$ rotates i in the $x-y$ plane; since it remains perpendicular to k it is rotated into the vector that was called e. It follows that $j=k\times i$ is rotated by $Z(\alpha)X(\beta)$ into the position of the vector f. It can now be seen that if $Z(\alpha)X(\beta)$ is preceded by $Z(\gamma)$ the result is $\mathbf{R}(\alpha,\beta,\gamma)$ since $Z(\gamma)$ leaves k unchanged but rotates the $x-y$ plane by γ; this means that i and j are rotated in the $x-y$ plane relative to e and f by an angle γ and, following $Z(\alpha)X(\beta)$, are in their final position.

The preceding argument shows that each rotation can be expressed by some choice of the Euler angles. The parametrization is, however, not unique for rotations about the z axis for which $\beta=0$, since any rotation $Z(\phi)$ can be expressed in the form $\mathbf{R}(\alpha, 0, \phi - \alpha)$ for arbitrary α. This indicates that the parametrization is singular at $\beta = 0$ and in particular at the identity. This fact complicates somewhat the problem of finding the invariant weight function on the group.

To conclude the discussion of the Euler angles we indicate how they may be determined for an arbitrary orthogonal matrix $\mathbf{R}$ with elements a_{ij} It can be seen by inspection of (9.6) that

$$\tan\alpha = -\frac{a_{13}}{a_{23}} \tag{9.8a}$$

$$\cos\beta = a_{33} \tag{9.8b}$$

$$\tan\alpha = \frac{a_{31}}{a_{32}} \tag{9.8c}$$

Equations (9.8) leave the quadrants of α and γ undetermined but these can be fixed from the signs of a_{13} and a_{31}.

Figure 9.4. *A rotation about v by ϕ constructed as successive rotations about u_1 and u_2 by π.*

In later applications the Euler angles A, B, Γ of the product $X(\beta)\, Z(\alpha)\, X(\beta')$ will be required. It can be verified by multiplying the matrices and comparing the result with (9.6) that they are given (implicitly) by

$$\sin A = \frac{\sin\alpha \sin\beta'}{\sin B} \tag{9.9a}$$

$$\cos A = \frac{\cos\beta\cos\alpha\sin\beta' + \sin\beta\cos\beta'}{\sin B} \tag{9.9b}$$

$$\cos B = \cos\beta\cos\beta' - \sin\beta\sin\beta'\cos\alpha \tag{9.9c}$$

$$\sin\Gamma = \frac{\sin\alpha\sin\beta}{\sin B} \tag{9.9d}$$

$$\cos\Gamma = \frac{\sin\beta\cos\alpha\cos\beta' + \cos\beta\sin\beta'}{\sin B} \tag{9.9e}$$

It is of interest to remark that any rotation can be expressed as the product of two rotations, each by π. To show this we suppose that a rotation is about an axis v by an angle ϕ. Let u_1 and u_2 be two vectors in the plane P perpendicular to v with an angle $\phi/2$ between them. A rotation by π about either u_1 or u_2 rotates P into itself and $\mathbf{v}$ into $-\mathbf{v}$. We consider a rotation about $\mathbf{u_1}$ by π followed by a rotation about $\mathbf{u_2}$ by π. The product of the two rotations is seen to leave $\mathbf{v}$ invariant and rotate P into itself. Furthermore, the first rotation leaves $\mathbf{u_1}$ invariant and the second rotation leaves $\mathbf{u_1}$ an angle $\phi/2$ on the other side of $\mathbf{u_2}$. The product of the rotations, therefore, rotates $\mathbf{u_1}$ (and all other vectors perpendicular to $\mathbf{v}$) through an angle ϕ. In Fig. 9-4 we show the axes $\mathbf{u_1}$ and $\mathbf{u_2}$ and a point that is rotated by ϕ in the plane perpendicular to $\mathbf{v}$.

The product of two rotations can be calculated from this fact. Consider rotations $\mathbf{R}_1$ and $\mathbf{R}_2$ by angles ϕ_1 and ϕ_2 about axes $\mathbf{v_1}$ and $\mathbf{v_2}$ respectively. We consider a vector $\mathbf{u}$ in the intersection of the planes P_1 and P_2 perpendicular to $\mathbf{v_1}$ and $\mathbf{v_2}$. Let $\mathbf{u_1}$ be a vector in P_1 such that the angle between $\mathbf{u_1}$ and $\mathbf{v_2}$ is $\phi_1/2$, and $\mathbf{u_2}$ be a vector in P_2 such that the angle between $\mathbf{u}$ and $\mathbf{u_2}$ is $\phi_2/2$. Then $\mathbf{R}_1$ can be constructed as a rotation about $\mathbf{u_1}$ by π followed by a rotation about $\mathbf{u_2}$ by π. Similarly, $\mathbf{R}_2$ is a rotation about $\mathbf{u}$ by π followed by a rotation about $\mathbf{u}$ by π. The product $\mathbf{R}_2\mathbf{R}_1$ ($\mathbf{R}_1$ is performed first) is then a product of four rotations by π; the second and third are both about $\mathbf{u}$ by π and, therefore, multiply to give $\mathbf{I}$. The product $\mathbf{R}_2\mathbf{R}_1$ is, therefore, a rotation about $\mathbf{u_1}$ by π followed by a rotation about $\mathbf{u_2}$ by π.

9.2 The functions v^α_β and invariant integration

In this section the functions v^α_β defined in equation (3.8) will be calculated for $O(3)+$. It is then possible to calculate the weight function for invariant integration immediately. It is unfortunately not useful to calculate the functions directly from (3.8) since the parametrization in the Euler angles is singular at e. It is necessary instead to introduce a new coordinate system into the neighborhood of e. To terms in first-order, an element in the neighborhood of e can be written in the form

$$B\left(h_1,h_2,h_3\right)=\begin{pmatrix} 1 & -h_3 & h_2 \\ h_3 & 1 & -h_1 \\ -h_2 & h_1 & 1 \end{pmatrix} \tag{9.10}$$

since, as we have seen, this is orthogonal in first-order in h.

For convenience we will replace α by α_1, β by α_2, and α by α_3. and denote by $\mathbf{h}$ and α vectors with components (h_1,h_2,h_3) and $(\alpha_1,\alpha_2,\alpha_3)$ respectively. We consider the product functions $\mathbf{f}\left(\mathbf{h},\alpha\right)$ that are defined to be the Euler angles of the product rotation $R\left(\alpha\right)B\left(h\right)$; that is,

$$v^r_c\left(\alpha\right)=\begin{pmatrix} \sin\gamma\,\csc\beta & \cos\gamma\,\csc\beta & 0 \\ \cos\gamma & -\sin\gamma\,\csc\beta & 0 \\ -\sin\gamma\,\cot\beta & -\cos\gamma\,\cot\beta & 1 \end{pmatrix} \tag{9.11}$$

where $v^r_c\left(\alpha\right)$ is the element in row r and column c.

The determinant of the matrix in (9.11) is observed to be $-\csc\beta$. The weight function for invariant group integration has been found to be the reciprocal of this determinant and ism therefore, $-\sin\beta$. The fact that the weight function is negative is rather disconcerting. This difficulty arises because the parametrization in the Euler angles is singular at e and there is no unique prescription for carrying the coordinates $\mathbf{h}$ into the coordinates α. Another manifestation of this difficulty is that $w(e)=0$. This shows that an attempt to evaluate w by (5.9) would necessarily fail since it was assumed in the derivation of (5.9) that $w(e)=1$. Since it is desirable that the weight function be positive we arbitrarily change the sign and write

$$w\left(\alpha,\beta,\gamma\right)=\sin\beta. \tag{9.12}$$

This sign change can be justified by permuting the rows of (1.3) since there is no *a priori* ordering of the Euler angles.

9.3 The homomorphism of $SU(2)$ onto $O(3)^+$

The Lie algebra of $O(3)$ is known to be composed of all skew-symmetric 3×3 matrices with real elements. A matrix of this form can be expressed as a linear combination of the matrices **i**, **j**, **k** of equation (4.50). Since these matrices satisfy the same com- mutation relations (up to a factor) as the matrices $\mathbf{I}_x$, $\mathbf{I}_y$, $\mathbf{I}_z$ of (8.7) the groups $O(3)$ and $SU(2)$ must be locally isomorphic. It will now be shown that there is a homomorphic mapping of $SU(2)$ onto the group $O(3)^+$.

We consider for each point in space with coordinates (x, y, z) the matrix

$$\mathbf{P}(x, y, z) = \begin{pmatrix} iz & ix - y \\ ix + y & -iz \end{pmatrix} \tag{9.13}$$

The matrix $\mathbf{P}(x, y, z)$ can be expressed in the form $xI_x + yI_y + zI_z$ where $\mathbf{I}_x$, $\mathbf{I}_y$, and $\mathbf{I}_z$ are the matrices defined in equation (8.7). It is observed that $\mathbf{P}$ is a skew-Hermitian matrix and has determinant $x^2 + y^2 + z^2$. It is also true that any 2×2 skew-Hermitian matrix with zero trace can be expressed in the form (9.13) for a suitable choice of x, y, z.

We consider now, for an arbitrary matrix $\mathbf{A} \in SU(2)$ the matrix $\mathbf{P}'$ defined by

$$\mathbf{P}' = \mathbf{A}\,\mathbf{P}'\,\mathbf{A}^{-1} = \mathbf{A}'\,\mathbf{P}'\,\mathbf{A}^{\dagger} \tag{9.14}$$

It is observed that $\mathbf{P}'^{\dagger} = \mathbf{A}\mathbf{P}^{\dagger}\mathbf{A}^{\dagger} = -\mathbf{A}\mathbf{P}\mathbf{A}^{\dagger} = -\mathbf{P}^{\dagger}$ so that $\mathbf{P}'$ is skew-Hermitian. Furthermore, $\mathrm{Tr}(\mathbf{P}') = \mathrm{Tr}(\mathbf{P}') = 0$ by the trace invariance property. It is, therefore, possible to write

$$\mathbf{P}' = \begin{pmatrix} iz' & ix' - y' \\ ix' + y' & -iz' \end{pmatrix}. \tag{9.15}$$

It follows from the form of (16) that the numbers x', y', z' are linear functions of x, y, z. It can, furthermore, be seen that the determinant of $\mathbf{P}'$ is equal to the determinant of $\mathbf{P}$, since $\left|\mathbf{A}\mathbf{P}\mathbf{A}^{-1}\right| = |\mathbf{A}|\,|\mathbf{P}|\,\left|\mathbf{A}^{-1}\right| = |\mathbf{P}|$. This result implies that

$$x'^2 + y'^2 + z'^2 = x^2 + y^2 + z^2 \tag{9.16}$$

and that the linear transformation on x, y, z generated by (9.14) is in fact a rotation. We conclude that for each $\mathbf{A} \in SU(2)$ there is a corresponding rotation $f(\mathbf{A}) \in 0(3)^+$.

The mapping f is a homomorphism since, if $\mathbf{A}_1$ and $\mathbf{A}_2$ are any two elements of $SU(2)$, $f(\mathbf{A}_1, \mathbf{A}_2)$ is the rotation generated by transforming $\mathbf{P}$ to

$$\mathbf{A}_1\mathbf{A}_2\mathbf{P}\left(\mathbf{A}_1\mathbf{A}_2\right)^{-1} = \mathbf{A}_1\left(\mathbf{A}_2\mathbf{P}\mathbf{A}_2\right)^{-1}\mathbf{A}_1^{-1}.$$

This is, however, the transformation generated by transforming $\mathbf{P}$ first by $\mathbf{A}_2$ and then by $\mathbf{A}_1$; the resulting rotation is that obtained by rotating first by $f(\mathbf{A}_2)$ and then by $f(\mathbf{A}_1)$ or $f(\mathbf{A}_1\mathbf{A}_2) = f(\mathbf{A}_1)f(\mathbf{A}_2)$.???

It will now be shown that each proper rotation is the image under f of the same element in $SU(2)$. This is proved by exhibiting explicitly elements of $SU(2)$

that generate rotations by an arbitrary angle about the x and z axes. We will require the following multiplication laws of the matrices $\mathbf{I}_x$, $\mathbf{I}_y$, $\mathbf{I}_z$.

$$\begin{aligned}\mathbf{I}_x\,\mathbf{I}_y &= -\mathbf{I}_y\,\mathbf{I}_x = \mathbf{I}_z\\ \mathbf{I}_y\,\mathbf{I}_z &= -\mathbf{I}_z\,\mathbf{I}_y = \mathbf{I}_x\\ \mathbf{I}_z\,\mathbf{I}_x &= -\mathbf{I}_x\,\mathbf{I}_z = \mathbf{I}_y.\end{aligned} \tag{9.17}$$

We consider now the matrix

$$\mathbf{A}\,(0,\phi,0) = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix} = \cos\phi + \sin\phi \mathbf{I}_z. \tag{9.18}$$

It is apparent that this matrix commutes with $\mathbf{I}_z$ so that

$$\mathbf{A}\,(0,\phi,0)\,\mathbf{P}\,\mathbf{A}\,(0,\phi,0)^\dagger = x'\mathbf{I}_x + y'\mathbf{I}_y + z'\mathbf{I}_z \tag{9.19}$$

Since z is unchanged, $\mathbf{A}\,(0,\phi,0)$ evidently generates a rotation about the z axis. We can calculate explicitly

$$\begin{aligned}&(\cos\phi + \sin\phi\mathbf{I}_z)\,(x\mathbf{I}_x + y\mathbf{I}_y + z\mathbf{I}_z)\,(\cos\phi - \sin\phi\mathbf{I}_z)\\ &\qquad = z\mathbf{I}_z + \left(\left(\cos^2\phi - \sin^2\phi\right)x - (2\sin\phi\cos\phi)\,y\right)\mathbf{I}_x\\ &\qquad\qquad + \left((2\sin\phi\cos\phi)\,x + \left(\cos^2\phi - \sin^2\phi\right)y\right)\mathbf{I}_y\\ &\qquad = (x\cos 2\phi - y\sin 2\phi)\,\mathbf{I}_x + (x\sin 2\phi - y\cos 2\phi)\,\mathbf{I}_y + z\mathbf{I}_z.\end{aligned}$$

This result indicates that

$$\begin{aligned}x' &= x\cos 2\phi - y\sin 2\phi,\\ y' &= x\sin 2\phi - y\cos 2\phi,\\ z' &= z.\end{aligned}$$

and that $\mathbf{A}\,(0,\phi,0)$ generates a rotation about the z axis by 2ϕ. It can be shown in the same way that the matrix

$$\mathbf{A}\,(\theta,0,0) = \begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix} = \cos\theta\,\mathbf{I} + \sin\theta\,\mathbf{I}_x \tag{9.20}$$

generates a rotation by 2θ about the x axis. Since an arbitrary proper rotation $R\,(\alpha,\beta,\gamma)$ can be expressed in the form $Z\,(\alpha)\,X\,(\beta)\,Z\,(\gamma)$, it can be generated by the element $\mathbf{A}\,(0,\alpha/2,0)\,\mathbf{A}\,(0,\beta/2,0)\,\mathbf{A}\,(0,\gamma/2,0)$ in $SU(2)$. This element can be written

$$\begin{aligned}&\begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix}\begin{pmatrix} \cos\frac{\beta}{2} & i\sin\frac{\beta}{2} \\ i\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix}\begin{pmatrix} e^{i\gamma/2} & 0 \\ 0 & e^{-i\gamma/2} \end{pmatrix}\\ &= \begin{pmatrix} e^{i(\alpha+\gamma)/2}\cos\frac{\beta}{2} & ie^{i(\alpha-\gamma)/2}\sin\frac{\beta}{2} \\ ie^{-i(\alpha-\gamma)/2}\sin\frac{\beta}{2} & e^{-i(\alpha+\gamma)/2}\cos\frac{\beta}{2} \end{pmatrix}\\ &= \mathbf{A}\left(\tfrac{\beta}{2}, \tfrac{\alpha+\gamma}{2}, \tfrac{\alpha-\gamma}{2}\right)\end{aligned} \tag{9.21}$$

It is important to observe that the mapping f of $SU(2)$ onto $0(3)^+$ is not an isomorphism; since $\mathbf{APA}^{-1} = (-\mathbf{A})\,\mathbf{P}\,(-\mathbf{A})^{-1}$ it is immediately apparent that $f(\mathbf{A}) = f(-\mathbf{A})$ and that the mapping f cannot be isomorphic. It is important to calculate the kernel of f, that is, the set of elements $\mathbf{A} \in SU(2)$ that satisfy $f(\mathbf{A}) = \mathbf{I}$. In order that $f(\mathbf{A}) = \mathbf{I}$, it is necessary that $\mathbf{A}$ satisfy $\mathbf{APA}^{-1} = \mathbf{P}$ or $\mathbf{AP} = \mathbf{PA}$ for all matrices $\mathbf{P}$ of the form (9.13). In particular, the matrices $\mathbf{I}_x$ and $\mathbf{I}_z$ must commute with $\mathbf{A}$. It can be observed from (8.6) and (8.7) that an arbitrary matrix $\mathbf{A}$ can be written

$$\mathbf{A} = \cos\theta\,\cos\phi + \cos\theta\,\sin\phi\,\mathbf{I}_z + \sin\theta\,\cos\psi\,\mathbf{I}_x + \sin\theta\,\sin\psi\,\mathbf{I}_y$$

In order that this commute with $\mathbf{I}_z$ it is necessary that

$$\sin\theta\,\cos\psi\,\mathbf{I}_y + \sin\theta\,\sin\psi\,\mathbf{I}_x = 0 \tag{9.22}$$

which implies $\sin\theta = 0$ and $\cos\theta = 1$. The requirement that $\cos\phi + \sin\phi\,\mathbf{I}_z$ commute with $\mathbf{I}_x$ implies that $\sin\phi = 0$ and $\cos\phi = \pm 1$. We can conclude that the kernel of the homomorphism consists of the matrices $\mathbf{I}$ and $-\mathbf{I}$. This subgroup is denoted by Z_2. To conclude this discussion, we can state that the group $0(3)^+$ is isomorphic to the factor group $SU(2)/Z_2$. The elements of this group are the cosets of Z_2 consisting of pairs of elements of $SU(2)$ which differ only in sign.

9.4 Representations of the rotation group

Certain of the representations $\mathbf{D}^j(\mathbf{A})_{mn}$ of $SU(2)$ given by equation (8.11) also provide representations of the group $0(3)^+$. In order that this occur it is sufficient that $\mathbf{D}^j(\mathbf{A}) = \mathbf{D}(-\mathbf{A})$; in this case each element of a coset of Z_2 is represented by the same matrix. The $\mathbf{D}^j(\mathbf{A})$, therefore, provide a representation of $SU(2)/Z_2$, and hence, also of $0(3)^+$. If the parameters of $\mathbf{A}$ are θ, ϕ, ψ the parameters of $-\mathbf{A}$ are θ, $\phi \pm \pi$, $\psi \pm \pi$. Inspection of (8.11) shows that $\mathbf{D}^j(\mathbf{A}) = \mathbf{D}(-\mathbf{A})$ if

$$e^{\pm i(m+n)\pi}\,e^{\pm i(n-m)\pi} = 1. \tag{9.23}$$

The left-hand side of (9.23) is one of $e^{\pm 2in\pi}$, $e^{\pm 2im\pi}$. If j is an integer m and n are each integers and (9.23) is satisfied; if $j = 1/2, 3/2, \ldots,$ $2m$ and $2n$ are each odd and $\mathbf{D}^j(-\mathbf{A}) = -\mathbf{D}^j(-\mathbf{A})$. It can be concluded that the matrices $\mathbf{D}^j(\mathbf{A})$ can be used to construct representations of $0(3)^+$ provided j is an integer.

The representations can be obtained immediately by substituting $\theta = \beta/2$, $\phi = (\alpha+\gamma)/2$, $\psi = (\alpha-\gamma)/2$ into (8.11). The result is

$$\begin{aligned}\mathbf{D}^j(\alpha,\beta,\gamma)_{mn} &= i^{m-n}e^{-im\alpha}e^{-in\gamma} \\ &\quad\times \sum_t (-1)^t\,\frac{((l+m)!\,(l-m)!\,(l+n)!\,(l-n)!)^{1/2}}{(l+m-t)!\,(t+n-m)!t!\,(l-n+t)!} \\ &\quad\times \cos^{2l+m-n-2t}\frac{\beta}{2}\,\sin^{2t+n-m}\frac{\beta}{2} \\ &= i^{m-n}e^{-im\alpha}d^l_{mn}(\beta)\,e^{-in\gamma}\end{aligned} \tag{9.24}$$

where

$$
\begin{aligned}
d^l_{mn}(\beta) = \times \sum_t (-1)^t \frac{((l+m)!\,(l-m)!\,(l+n)!\,(l-n)!)^{1/2}}{(l+m-t)!\,(t+n-m)!t!\,(l-n+t)!} \\
\times \cos^{2l+m-n-2t}\frac{\beta}{2}\sin^{2t+n-m}\frac{\beta}{2}.
\end{aligned}
\tag{9.25}
$$

It should be pointed out that we have used the symbol $\mathbf{D}$ to denote two different functions defined by (9.24) and (8.11). This should, however, give rise to no confusion.

The representations $D^l(R)$ are unitary and irreducible since they are unitary and irreducible as representations of $SU(2)$. They also exhaust the irreducible representations of $0(3)^+$, since any other irreducible representation would give rise to an irreducible representation of $SU(2)$ with the property that $\mathbf{D}(\mathbf{A}) = \mathbf{D}(-\mathbf{A})$. It is known, however, that the only irreducible representations with this property are the $\mathbf{D}^j(\mathbf{A})$, j an integer.

It can be seen from (8.13) and (8.20) – (8.22) that, in the special cases $m = \pm l$, or $n = \pm l$, the functions $d^l_{mn}(\beta)$ are given by

$$
\begin{aligned}
d^l_{-l,n}(\beta) &= \sqrt{\frac{2l}{l+n}}\cos^{l-n}\frac{\beta}{2}\sin^{l+n}\frac{\beta}{2} \\
d^l_{l,n}(\beta) &= (-1)^{l-n}\sqrt{\frac{2l}{l+n}}\cos^{l+n}\frac{\beta}{2}\sin^{l-n}\frac{\beta}{2} \\
d^l_{m,-l}(\beta) &= (-1)^{l+m}\sqrt{\frac{2l}{l+m}}\cos^{l-m}\frac{\beta}{2}\sin^{l+m}\frac{\beta}{2} \\
d^l_{m,l}(\beta) &= \sqrt{\frac{2l}{l+m}}\cos^{l+m}\frac{\beta}{2}\sin^{l-m}\frac{\beta}{2}
\end{aligned}
\tag{9.26}
$$

The representation $\mathbf{D}^1(\alpha,\beta,\gamma)$ can be calculated to be

$$
\frac{1}{2}\begin{pmatrix}
e^{i(\alpha+\gamma)}(1+\cos\beta) & -i\sqrt{2}\,e^{i\alpha}\sin\beta & e^{i(\alpha-\gamma)}(1-\cos\beta) \\
-i\sqrt{2}e^{i\gamma}\sin\beta & \cos\beta & -i\sqrt{2}e^{-i\gamma}\sin\beta \\
-e^{i(\gamma-\alpha)}(1-\cos\beta) & -i\sqrt{2}e^{-i\alpha}\sin\beta & e^{i(\alpha+\gamma)}(1+\cos\beta)
\end{pmatrix}
\tag{9.27}
$$

It can be observed from (9.27) that rotations about the z axis are represented by diagonal matrices. This is in general the case; it follows immediately from (8.12) that

$$
\mathbf{D}^l(\alpha,0,\gamma)_{mn} = e^{-i(\alpha+\gamma)m}\delta_{mn}
\tag{9.28}
$$

The group element $\mathbf{R}(\alpha,0,\gamma)$ is, of course, a rotation about the z axis by $\alpha+\gamma$. Rotations about the x axis are represented by

$$
\mathbf{D}^l(0,\beta,0)_{mn} = i^{m-n}d^l_{mn}(\beta)
\tag{9.29}
$$

It is not difficult to see that a rotation about the y axis by an angle β can be generated by rotating first about the z axis by $-\pi/2$, rotating about the x axis by

β and then rotating about the z axis by $\pi/2$. The rotation is, therefore, represented by

$$\mathbf{D}^l\left(\frac{\pi}{2}, \beta, -\frac{\pi}{2}\right)_{mn} = d^l_{mn}(\beta) \tag{9.30}$$

The representation (9.24) differs from that frequently given in that it includes an additional factor of i^{m-n}. This has been included to compensate for the fact that the Euler angles have been defined in the classical way so that the second rotation is about the x axis rather than about the y axis as is the case in most quantum-mechanical applications. The present results relations can be transcribed to the usual quantum-mechanical phase conventions by deleting the factor i^{m-n}, and regarding a rotation about the x axis as being about the y axis and a rotation about the y axis as a negative rotation about the x axis. The phases of the spherical harmonics, to be discussed in the next section, conform to the usual convention because of the inclusion of the extra factor i^{m-n}.

9.5 Harmonic polynomials and representations of $O(3)^+$

Another possible method of constructing representations of the rotation group is to consider homogeneous polynomials of fixed degree in the variables x, y, and z. The methods described in section §(**??**) can be applied to construct representations in the invariant subspaces of such functions. This method, which was used successfully in section §(**??**), is less satisfactory for the present problem since it does not generate irreducible representations. This method will, however, be discussed in this section to demonstrate the relation between the group representations obtained in the previous section and the important functions, the spherical harmonics.

We consider the space S_l of homogeneous polynomials of degree l in the variables x, y, z. This space is spanned by the monomials of the form $x^m y^n z^{l-m-n}$. There are

$$\sum_{m=0}^{l}\sum_{m=0}^{l-m} = \sum_{m=0}^{l}(l-m+1) = \frac{(l+2)(l+1)}{2}$$

such monomials, which are clearly linearly independent, so that S_l is of dimension $(l+2)(l+1)/2$. The representations defined by S_l must be reducible since S_l contains an invariant subspace, the space S_{l-2} of all polynomials of the form $(x^2+y^2+z^2)P_{l-2}$ where P_{l-2} is a polynomial of degree $l-2$. Since the group is orthogonal, the function $x^2+y^2+z^2$ is invariant under the group transformations and S_{l-2} is an invariant subspace (of dimension $l(l-1)/2$). The representation defined by S_l can be assumed to be unitary, in which case the subspace $S^{\perp}_{l-2}$ orthogonal to S_{l-2} is also invariant. This subspace is of dimension

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

This invariant subspace is rather nebulous since no inner product has been defined on S_l. It is Possible, however, to construct a $(2l+1)$−dimensional invariant subspace of S_l in another way. We consider the mapping of S_l onto S_{l-2} defined by

$$P_l(x) \to P_{l-2}(x) = \nabla^2 P_l(x)$$

where the image of P_l is obviously in S_{l-2}. It is convenient to denote the points whose coordinates are in (x, y, z) by $\mathbf{x}$. It will now be shown that, for any rotation R,

$$\nabla^2 \left(P_l\left(\mathbf{R}^{-1}x\right)\right) = P_{l-2}\left(\mathbf{R}^{-1}x\right)$$

where $P_{l-2}(x) = \nabla^2 P_l(x)$. We will denote $\mathbf{R}^{-1}\mathbf{x}$ by $\mathbf{x}'$. If the elements of $\mathbf{R}$ are a_{ij} the component j of $\mathbf{x}'$ is given by $\sum_i a_{ij}x_i$. We can now write

$$\begin{aligned}
\sum_i \frac{\partial^2}{\partial \mathbf{x}_i^2} P_l\left(\mathbf{R}^{-1}\mathbf{x}\right) &= \sum_{ijk} \frac{\partial^2 P_l}{\partial \mathbf{x}'_j \mathbf{x}'_k}\left(\mathbf{x}'\right) \frac{\partial \mathbf{x}'_j}{\partial \mathbf{x}_i} \frac{\partial \mathbf{x}'_k}{\partial \mathbf{x}_i} \\
&= \sum_{ijk} a_{ij} a_{ik} \frac{\partial^2 P_l}{\partial \mathbf{x}'_j \mathbf{x}'_k}\left(\mathbf{x}'\right) \\
&= \sum_j \frac{\partial^2 P_l}{\partial \mathbf{x}'^2_j}\left(\mathbf{x}'\right) \\
&= P_{l-2}\left(\mathbf{R}^{-1}\mathbf{x}\right)
\end{aligned}$$

The implication of this result is that the mapping ∇^2 from S_l onto S_{l-2} satisfies

$$\nabla^2 D^l(\mathbf{R}) = D^{l-2}(\mathbf{R})\nabla^2 \tag{9.31}$$

where $D^l(\mathbf{R})$ are defined by equation (**??**).

We consider now the subspace H_l of S_l composed of polynomials P_l satisfying

$$\nabla^2 P_l = 0. \tag{9.32}$$

A function satisfying this equation, which is Laplace's equation, is said to be **harmonic**. It follows immediately from (9.31) that H_l is invariant; if $\nabla^2 P_l = 0$,

$$\nabla^2 D^l(\mathbf{R}) P_l = D^{l-2}(\mathbf{R}) \nabla^2 P_l = 0$$

and $D^l(\mathbf{R}) P_l$ is also harmonic. It will now be shown that the representation of the rotation group generated by H_l, the set of harmonic polynomials of degree l, is equivalent to the representation $D^l(\mathbf{R})$ defined by (**??**).

It is possible to obtain $2l+1$ linearly independent solutions of (9.32) explicitly. For this purpose it is convenient to introduce new variables

$$\begin{aligned}
u &= \frac{1}{2}(x + iy), \\
v &= \frac{1}{2}(x - iy)
\end{aligned}$$

in terms of which (9.32) becomes

$$\frac{\partial^2 P_l}{\partial u \partial v} + \frac{\partial^2 P_l}{\partial z^2} = 0. \tag{9.33}$$

If P_l is a homogeneous polynomial of degree l in u, v, and z it is also a homogeneous polynomial of degree l in x, y, and z. It is possible to write down four solutions of (9.33), u^l, v^l, $u^{l-1}z$, $v^{l-1}z$ immediately. More generally, we look for a solution that contains a term of the form $u^{l-m}v^m$, $m = 0, 1, \ldots l$. This is not a solution since

$$\nabla^2 u^{l-m}v^m = (l-m)\, m u^{l-m-1}v^{m-1}. \tag{9.34}$$

It is possible to eliminate the right-hand side by adding to $u^{l-m}v^m$ a term $(-1)(l-m)mu^{l-m-1}v^{m-1}z^2/2$. One then obtains

$$\nabla^2 \left(u^{l-m}v^m - \frac{(l-m)mu^{l-m-1}v^{m-1}z^2}{2} \right) \\ = -\frac{(l-m)(l-m-1)m(m-1)u^{l-m-2}v^{m-2}z^2}{2}$$

It is now possible to add a third term, $(l-m)(l-m-1)m(m-1)u^{l-m-2}v^{m-2}/4!$ to eliminate the new term on the right-hand side. Proceeding in this way one eventually obtains a harmonic polynomial of degree l which can be written

$$f_{lm}(u, v, z) = \sum_p (-1)^p \frac{(l-m)!m!}{(l-m-p)!(m-p)!(2p)!} u^{l-m-p}v^{m-p}z^{2p} \tag{9.35}$$

The sum on p is from 0 to the smaller of m and $l-m$. It can be verified by direct substitution into (9.33) that f_{lm} is a harmonic polynomial. The functions f_{lm} have the further property, which will prove to be important, that the difference of the exponents of u and v, $l-2m$, is the same for each term. We note that there are $(l+1)$ functions f_{lm}.

In a similar way, it is possible to find solutions that contain a term $u^{l-m}v^{m-1}z$, where $m = 1, 2, \ldots, l$. These solutions can be written

$$g_{lm} \;=\; \sum_p (-1)^p \frac{(l-m)!(m-1)!}{(l-m-p)!(m-p-1)!(2p+1)!} u^{l-m-p}v^{m-p-1}z^{2p+1} \tag{9.36}$$

In this case the index p runs from 0 to the smaller of $m-1$ and $l-m$. There are l such solutions with the property that the difference of the exponents of u and v is $l-2m+1$ for each term in the sum.

There are altogether $(2l+1)$ functions f_{lm}, g_{lm}. These can be labeled by an index s,

$$s = -l, -l+1, \ldots\ldots, l-1, l,$$

s being the difference of the exponents of u and v in each term of a particular function. These functions will be denoted by k_{ls}. The functions k_{ls} are obviously linearly independent since no two of them can contain the same monomial. It can also be seen that the functions k_{ls} span H_l. Let $P_l(x)$ be a harmonic polynomial of degree l. Consider a term $u^{p-m}v^{q-m}z^{l-p-q+2m}$ in P_l; it can be seen from (9.33) that the coefficient of every term in P_l of the form $u^pv^qz^{l-p-q}$ is uniquely determined by the coefficient of $u^pv^qz^{l-p-q}$. In fact, all the terms of this form must occur as a constant multiple of $k_{l,p-q}$, and can be removed by subtracting $ck_{l,p-q}$ from P_l for some c. It is, therefore, apparent that P_l can be expressed as a linear combination of the k_{ls}.

The functions k_{ls} generate, by equation (**??**), a representation of the proper rotation group. This representation will be denoted by $\Delta^l(R)$. We will not calculate $\Delta^l(R)$ explicitly but rather show that it is equivalent to the representation $D^l(R)$ defined by (**??**). It will be shown first that $\Delta^l(R)$ is diagonal if R is a rotation by an angle ϕ about the z axis. Under the inverse of such a rotation, the spatial variables are transformed according to

$$\begin{aligned} x &\to \quad x\cos\phi + y\sin\phi, \\ y &\to -x\sin\phi + y\cos\phi, \\ z &\to z. \end{aligned}$$

It follows that u is transformed according to

$$u \to (x\cos\phi + y\sin\phi) + i(-x\sin\phi + y\cos\phi) = e^{-i\phi}x + ie^{-i\phi}y = e^{-i\phi}u.$$

Similarly, $v = u^*$ is transformed to $e^{i\phi}v$. It follows that a monomial of the form $u^\alpha v^\beta z^\gamma$ is transformed to $e^{i(\beta-\alpha)\phi}u^\alpha v^\beta z^\gamma$, and hence that

$$k_{ls}\left(R^{-1}x\right) = c^{-is\phi}k_{ls}(x)$$

since each term of k_{ls} is multiplied by the same factor $e^{is\phi}$ From (**??**) we can write, for R a rotation by ϕ about the z axis,

$$\Delta^l(R)_{st} = e^{is\phi}\delta_{st} \tag{9.37}$$

Each class of the rotation group has been shown to contain a rotation about the z axis. Comparison of (**??**) and (9.37) shows that rotations about the z axis are represented by the same matrices in D and Δ; the characters of the two representations are, therefore, the same and the representations are equivalent.

We consider now the matrix M that transforms $D^l(R)$ to $\Delta^l(R)$:

$$M^{-1}\Delta^l(R)M = D^{l(R)} \tag{9.38}$$

for all R. If R is, in particular, a rotation about the z axis, $\Delta^l(R) = D^l(R)$ and M satisfies $MD^l(R) = D^l(R)M$. In this case $D^l(R)$ is, however, diagonal with diagonal elements which are in general different. It has been seen previously that this implies M is diagonal; the matrix elements of M will, therefore, be denoted by $\mu_i\delta_{ij}$. Equation (9.38) can now be written

$$\Delta^l(R)_{mn} = \mu_m D^l(R)_{mn}\mu_n^{-1}. \tag{9.39}$$

The functions $k_{ls}(x)$ and the representations $\Delta^l(R)$ are related by

$$k_{lt}\left(R^{-1}x\right) = \sum_s \Delta^l(R)_{st}k_{ls}(x).$$

Substituting (9.39) into this relation yields the result

$$\mu_t k_{lt}\left(R^{-1}x\right)\sum_s D^l(R)_{st}\mu_s k_{ls}(x). \tag{9.40}$$

It can be concluded that the representation $D^l(R)_{st}$ that was obtained in equation (**??**) is also the representation generated by the harmonic polynomials $\mu_s k_{ls}(x)$. We will, henceforth, consider these functions rather than the functions f and g defined in equations (9.35) and (9.36), from which they differ by an undetermined factor.

The coordinates of the point x can be expressed in spherical polar coordinates as $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$. If the functions $\mu_m k_{lm}(x)$ are written in terms of these coordinates it is evident that they have the form $r^l Y_{lm}(\theta,\phi)$ where $Y_{lm}(\theta,\phi)$ is a polynomial in $\sin\theta$, $\cos\theta$, $\sin\phi$, and $\cos\phi$. The functions $Y_{lm}(\theta,\phi)$ are the important **spherical harmonics**. Since $r^l Y_{lm}(\theta,\phi)$ must satisfy Laplace's equation in spherical polar coordinates, the spherical harmonics must satisfy

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial Y_{lm}}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2 Y_{lm}}{\partial\phi}+l(l+1)Y_{lm}=0. \tag{9.41}$$

It will now be shown that for fixed l the spherical harmonics are determined up to an arbitrary constant by equation (9.40). We denote by (θ,ϕ) and (θ',ϕ') the angular coordinates of the points x and $R^{-1}x$ respectively. In terms of the spherical harmonics (9.40) becomes

$$Y_{lm}(\theta',\phi')=\sum_n D^l(R)_{nm}Y_{ln}(\theta,\phi). \tag{9.42}$$

If $\mathbf{R}$ is a rotation about the z axis by a the angles (θ',ϕ') are simply $(\theta,\phi-\alpha)$. USing equation (**??**), we can write (9.42) as

$$Y_{lm}(\theta,\phi-\alpha)=e^{im\alpha}Y_{lm}(\theta,\phi).$$

Putting $\phi=0$ and changing the sign of α, we obtain

$$Y_{lm}(\theta,\alpha)=e^{im\alpha}Y_{lm}(\theta,0), \tag{9.43}$$

indicating that the only dependence of Y_{lm} on the azimuthal angle is in the factor $e^{im\alpha}$.

In the direction of the positive z axis the spherical harmonics must be independent of the azimuthal angle, that is, $Y_{lm}(0,\alpha)=Y_{lm}(0,0)$. If $m\neq 0$, however, this can only be the case if $Y_{lm}(0,0)=0$ in view of the known dependence of Y_{lm} on the azimuthal angle. On the other hand $Y_{l0}(0,0)\neq 0$ since otherwise the spherical harmonics would vanish identically. The value of $Y_{l0}(0,0)$ will be chosen arbitrarily to be 1 so that

$$Y_{lm}(0,\phi)=\delta_{m0}. \tag{9.44}$$

We now put $\theta=\phi==0$ in (9.42) and replace $\mathbf{R}$ by $\mathbf{R}^{-1}$; the result is, since $\mathbf{D}^l\left(\mathbf{R}^{-1}\right)_{nm}=\mathbf{D}^l\left(\mathbf{R}\right)^*_{nm}$,

$$Y_{lm}(\theta',\phi')=\mathbf{D}^l\left(\mathbf{R}\right)^*_{m0}.$$

where θ', ϕ' are the polar angles of the direction into which $\mathbf{R}$ rotates the z axis. The rotation $\mathbf{Z}\,(\alpha + \pi/2)\,\mathbf{X}\,(\beta)$ is known from the definition of the Euler angles to rotate the vector k parallel to the z axis into the direction whose angular coordinates are (β, α). We can, therefore, write,

$$Y_{lm}\,(\beta, \alpha) = \mathbf{D}^l \left(\alpha + \frac{\pi}{2}, \beta, 0\right)^*_{m0} \tag{9.45}$$

or

$$Y_{lm}\,(\theta, \phi) = e^{im\phi} d^l_{m0}(\theta). \tag{9.46}$$

This result will be applied to obtain various properties of the spherical harmonics.

9.6 Differential equations for the group representations

It was shown in Section §(**??**) that the representations of a Lie group must satisfy the partial differential equations (**??**). In this section we will obtain these equations for the representations $\mathbf{D}'^l(R)$ of the rotation group. This task is again complicated by the singularity in the Euler angle coordinate system at the identity, so that