M. E. Rose

Elementary Theory of

Angular Momentum

ELEMENTARY THEORY OF ANGULAR MOMENTUM

M. E. ROSE

Chief Physicist
Oak Ridge National Laboratory

New York · JOHN WILEY & SONS, Inc.
London · CHAPMAN & HALL, Ltd.
1957

Library of Congress Catalog Card Number: 57–8893

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

During the past several years the theory of angular momentum has come to occupy a more and more important position in the development of physical theories of nuclear and atomic structure. One reason for this is found in those improvements in experimental techniques whereby it has become possible to measure angular distributions in nuclear reactions, or alternatively angular correlations of successively emitted radiations. From an entirely different quarter we find that the theory, in its modern form, can be very advantageously applied to the formulation and solution of problems associated with the static magnetic and electric nuclear moments which are coupled to the electric and magnetic fields arising from surrounding charges. Here reference is made to problems encountered in low-temperature studies as well as to those met in microwave spectroscopy.

It is therefore hardly surprising that there has been an upsurge of interest in the elements of the theory which are basic for the description of such physical phenomena. The theory of angular momentum is essentially a highly formal one. Its principal ingredients are certain parts of group theory and tensor algebra. In these pages, however, the emphasis is much less abstract than these formidable terms imply. The present work is the result of a course of lectures given at the Oak Ridge National Laboratory in the winter and spring of 1955. In presenting those lectures and in writing these pages it has been my conviction that a clear understanding of the elements of the theory could be helpful to many physicists, and that the ideas as well as the techniques involved can be made available to a majority of those with a graduate-course knowledge of quantum mechanics. There is no implication that one can avoid the use of formalism. The simplification which the present treatment, it is hoped, does achieve is based on two delimiting factors. First, we are concerned here only with the properties of rotations because of their intimate connection with the concept of angular momentum. Second, the reasoning is inductive, and, as the theory initially develops, it makes a "smooth-join" with those aspects of quantum mechanics that are, comparatively speaking, common knowledge. In this way, it is felt that it is possible to make the ideas as well as the analysis transparent and simple. At the same time, this entails little if any essential loss in power and/or elegance in the methods. What is sacrificed is the opportunity to discuss and treat the most general and complicated problems in the most expeditious manner. Actually, this loss of generality is felt in very few places,

and these are pointed out in the text. What is gained, it is hoped, is the opportunity to make the ideas underlying the theory and the means of using it available to a much wider audience.

Of course, a number of applications are given as illustrations. No attempt is made to treat all possible applications, nor is any one treated in anything resembling an exhaustive manner. In several cases this would hardly be necessary. For instance, we find that the use of the theory greatly facilitates the calculation of the matrix elements of the quadrupole coupling energy. Once this is done, the determination of the energy levels of the system under consideration is a matter of solving a secular determinant, and further discussion would, therefore, be irrelevant from the point of view of our expressed purpose. Part A is devoted to an exposition of the major elements of the theory. Most, though not all, of the applications are to be found in Part B.

The first chapter is a review of basic principles. Here we present such discussion of operators and their matrix representations as is pertinent for the sequel. In this connection a word of explanation or apology may be in order. To make the discussion more apposite a few references to physical situations are made, and the expression "angular momentum" or "spin" is used. While at this point no formal definitions have been given, it is felt that the reader has some "intuitive" idea of what is meant and would profit by the examples chosen for illustration and the discussion of them. For the purist, these examples would be meaningful on rereading after covering material appearing in Chapter II. Actually, the examples chosen are quite familiar ones, appearing in fairly elementary discussions. The remaining four chapters comprising Part A carry the development through the introduction of the coupling coefficients for vector addition (C-coefficients), the transformation properties of the angular momentum wave functions under rotations of the coordinate axes, irreducible tensors, and Racah coefficients. The applications presented in six chapters of Part B deal with static moments of systems composed of charged particles and elementary magnetic dipoles, particles of intrinsic spin $\frac{1}{2}$ and 1, oriented nuclei (this topic includes angular correlation in cascade disintegrations as well as angular distributions and changes in total cross sections in low-temperature experiments), coupling schemes in nuclear reactions, and wave functions for systems of identical particles.

It is a pleasure to record my gratitude to Drs. L. D. Roberts and J. W. T. Dabbs, for many conversations in which they ably represented the reader's point of view. Dr. A. E. Glassgold rendered helpful service in preparing a first draft of the notes for Part A.

M. E. Rose

April 1957

CONTENTS

Part A. GENERAL THEORY

Part B. APPLICATIONS

PART A

General Theory

I. REVIEW OF BASIC PRINCIPLES

The background needed for the ensuing chapters is presumably well known to all those who have an acquaintance with quantum mechanics. For ease of reference and for completeness, this chapter provides a brief résumé of the necessary formal constructs.

1. HERMITIAN OPERATORS

To discuss the properties of Hermitian operators we assume that the scalar product of two functions, χ and ζ for example, is defined and has these properties:

$$(\chi, \zeta) = (\zeta, \chi)^* \tag{1.1a}$$

$$(\chi, c\zeta) = c(\chi, \zeta) \tag{1.1b}$$

$$(\chi, \zeta_1 + \zeta_2) = (\chi, \zeta_1) + (\chi, \zeta_2) \tag{1.1c}$$

$$(\chi_1 + \chi_2, \zeta) = (\chi_1, \zeta) + (\chi_2, \zeta) \tag{1.1d}$$

The asterisk indicates complex conjugate; c is a complex number. The detailed form of the scalar product is discussed below.

The Hermitian adjoint Ω^+ of the operator Ω has the property

$$(\Omega^+\chi, \zeta) = (\chi, \Omega\zeta) \tag{1.2}$$

For example, for the gradient operator $\nabla^+ = -\nabla$ if the functions χ and ζ vanish on the boundary of the domain of integration. A Hermitian or self-adjoint operator is its own adjoint. This means that $\Omega = \Omega^+$, and therefore

$$(\Omega\chi, \zeta) = (\chi, \Omega\zeta) \tag{1.3}$$

In the example given above $i\nabla$ is Hermitian.

If we consider the matrix representation of the operator Ω, its Hermitian adjoint is the matrix obtained by transposing (interchanging rows and columns) and taking the complex conjugate of each element. Equation (1.2) may be regarded as the formal expression of this definition since, in conjunction with (1.1a), it shows that $(\Omega^+\chi, \zeta) = (\Omega\zeta, \chi)^*$ or $(\chi, \Omega\zeta) = (\zeta, \Omega^+\chi)^*$. Equation (1.3) is then the formal definition of a

Hermitian matrix. Whether we think of an operator or its matrix representation it follows easily that

$$(\Omega_1\Omega_2\Omega_3 \cdots \Omega_n)^+ = \Omega_n^+ \cdots \Omega_3^+\Omega_2^+\Omega_1^+$$

Two important properties of Hermitian operators are the following:
(*a*) They have real expectation values [from (1.3) and (1.1*a*)].

$$(\chi, \Omega\chi) = (\chi, \Omega\chi)^* \tag{1.4}$$

(*b*) The eigenvalue problem yields real eigenvalues and orthogonal functions. That is, the equation

$$\Omega\psi = \omega\psi$$

with its boundary conditions is satisfied only for certain values of ω, in this case, from (1.4) with $\chi = \psi$, real eigenvalues.[1] The corresponding eigenfunctions are labeled with the eigenvalues ω

$$\Omega\psi_\omega = \omega\psi_\omega \tag{1.5}$$

Consider two linearly independent eigenfunctions ψ_ω and $\psi_{\omega'}$. By linearly independent we mean there exist no non-vanishing constants c and c' such that $c\psi_\omega + c'\psi_{\omega'} = 0$. Otherwise ψ_ω and $\psi_{\omega'}$ would differ only by a multiplicative constant and would be equivalent.[2] From the definition (1.3) of a Hermitian operator it follows that

$$(\psi_{\omega'}, \Omega\psi_\omega) = (\Omega\psi_{\omega'}, \psi_\omega)$$

or

$$(\omega' - \omega)(\psi_{\omega'}, \psi_\omega) = 0$$

If $\omega \neq \omega'$, ψ_ω and $\psi_{\omega'}$ are orthogonal; that is, $(\psi_{\omega'}, \psi_\omega) = 0$. If $\omega = \omega'$, ψ_ω and $\psi_{\omega'}$ may still be orthogonalized since they are linearly independent. Thus, the linear combination $\psi_\omega - (\psi_{\omega'}, \psi_\omega)\psi_{\omega'}$ is orthogonal to $\psi_{\omega'}$ although these two functions have the same eigenvalue—are still degenerate. Actually these eigenfunctions have to be simultaneous eigenfunctions of a whole set of commuting Hermitian operators,[3] and the label ω stands for a set of quantum numbers. Then, in the comparison of independent eigenfunctions of the system, no pair will have identical sets of quantum numbers or eigenvalues. The degeneracy means simply that the eigenvalues of at least one of the operators will coincide. In the usual terminology degeneracy implies equal energy.

[1] We shall restrict ourselves to cases where the eigenvalue spectrum is complete.

[2] If we require that they have the same normalization they would, in fact, differ by only a phase factor (a number of modulus one).

[3] More will be said about the simultaneous diagonalization of a set of commuting Hermitian operators in section 3.

In view of the foregoing we can conclude that the scalar product of two independent eigenfunctions is

$$(\psi_{n'}, \psi_n) = 0 \quad \text{for} \quad n' \neq n$$

Here it is again emphasized that n and n' are collective labels for all the eigenvalues; $n' \neq n$ means that at least one of the eigenvalues is different for the two eigenfunctions. The eigenfunctions can be normalized to unity by multiplication by an appropriate number,

$$(\psi_{n'}, \psi_n) = 1 \quad \text{for} \quad n' = n$$

Thus, the ψ_n are said to constitute an orthonormal set

$$(\psi_{n'}, \psi_n) = \delta_{n'n} \tag{1.6}$$

Furthermore, they are also a complete set, although we shall not prove this here. By completeness we mean there exists no function F such that $(\psi_n, F) = 0$ for all n. As a consequence, a reasonably behaved [4] function f can be expanded in terms of the basis ψ_n:

$$f = \sum_n f_n \psi_n, \qquad f_n = (\psi_n, f) \tag{1.7}$$

The geometrical interpretation of the foregoing is obtained by regarding the set of ψ_n as a set of orthonormal unit vectors in a space of as many dimensions as the set of numbers n represents. This number need not be finite, of course. The expansion given above is the analogue of the representation of "vector" f in terms of its components f_n and the unit vectors ψ_n. Even when the space of n is not finite we are sometimes interested in a subspace which is. Thus, n may represent three numbers n_1, n_2, n_3. Of these, n_1 (the energy, say) has a spectrum with an infinite number (countable or otherwise) of permitted values, n_2 and/or n_3 may represent eigenvalues of Hermitian operators for which the spectrum involves a finite number of permitted values. An example is one of the components of the angular momentum. If n_3 has this property, then the sum (1.7) has a finite number of terms, with n_1 and n_2 fixed. This corresponds to choosing only those "vectors" f that lie in a subspace of the total space.

If the functions under discussion depend only on the space coordinate x, the scalar product in (1.1) is an integral

$$(\chi, \zeta) = \int d^3x \, \chi^*(x) \, \zeta(x)$$

over all values of x. Here d^3x is the volume element. In addition, we will be interested in eigenfunctions which also depend on variables with

[4] For those insisting on mathematical rigor a satisfactory definition of "reasonably behaved" is difficult. One can say that the coefficients f_n must exist and that the series (1.7) be summable in some sense, although this is tautological. Nevertheless, it may suffice to say that no difficulties are anticipated in our applications.

discrete ranges. For example, there may be a spinor index s. As an example (which will become clearer later on) the index s assumes $2j+1$ values for a particle of angular momentum [5] j. The wave function can then be written as a column matrix each element of which is a function of the space coordinates,

$$\chi(x) = \begin{bmatrix} \chi^{(1)}(x) \\ \chi^{(2)}(x) \\ \cdot \\ \cdot \\ \cdot \\ \chi^{(2j+1)}(x) \end{bmatrix}$$

and the scalar product involves a sum as well as an integration

$$(\chi, \zeta) = \sum_m \int d^3x \, \chi^{(m)*}(x) \, \zeta^{(m)}(x) = \int d^3x \, \chi^+ \zeta$$

In the last equality matrix multiplication of χ^+ into ζ is implied, and here the cross means complex conjugate transpose. Unless otherwise indicated, a scalar product involves summation over all independent variables, which will be an integration over variables with a continuous range and the usual sum for discrete variables.

As already indicated, the sum in the scalar product written above can be considered as the matrix multiplication of the row matrix for χ, obtained by transposing and complex conjugating (i.e., χ^+), times the column matrix for ζ. We can also introduce in this connection the orthonormal set ϕ_m whose elements are all zero except that one in the mth place which is unity. If there are $2j+1$ rows, so that j is a half-integer or an integer, then the ϕ_m look like this:

$$\phi_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \qquad \phi_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \end{bmatrix} \qquad \phi_{2j+1} = \begin{bmatrix} \\ \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

[5] Angular momentum will be expressed in units of $\hbar$. Thus j and similar symbols are angular momenta divided by $\hbar$. Stated otherwise, we depart from the ordinary cgs units and adopt a new system of mechanical units in which $\hbar$ has the numerical value unity.

Consequently, we can write

$$\chi = \sum_m \chi^{(m)} \phi_m$$

This checks with the foregoing when we realize that the ϕ_m form an orthonormal set. That is,

$$(\phi_m, \phi_n) = \delta_{nm}$$

The case $j = \frac{1}{2}$ is well known from the Pauli treatment of electron spin; here m assumes just two values, which we can take to be $\pm\frac{1}{2}$ for reasons to appear in section 8. Another possibility is intrinsic spin 1, of which the Maxwell field affords an excellent example. A sufficient and complete description for monochromatic fields is given in terms of the vector potential [6] **A** or its components which may be represented as follows:

$$A_1 = -\frac{1}{\sqrt{2}}(A_x + iA_y)$$

$$A_0 = A_z$$

$$A_{-1} = \frac{1}{\sqrt{2}}(A_x - iA_y)$$

The reasons for the phases and normalization will be discussed later. The "three-ness" $(2 \cdot 1 + 1)$ of this field arises from its vector character, which makes the intrinsic spin 1; see Chapters V and VII. Of course, as we shall see, the electromagnetic field can have any integer angular momentum ≥ 1, just as an electron can have any half-integer angular momentum $\geq \frac{1}{2}$.

2. UNITARY TRANSFORMATIONS

Of particular importance in the theory of angular momentum are unitary transformations. A unitary transformation is a linear homogeneous transformation which preserves lengths and angles; that is, scalar products are left invariant. The geometrical interpretation is a rotation in the space spanned by the base vectors ψ_n discussed in the last section. That is, we now choose a different basis, φ_n say, and, by completeness of the set ψ_n, the functions φ_n must be expressible as linear combinations

[6] We assume that the vector potential is defined with the aid of the Lorentz condition $\nabla \cdot \mathbf{A} + (1/c)(\partial\phi/\partial t) = 0$, where ϕ = scalar potential. Then the scalar potential is defined by **A** to within an additive time-independent quantity. The latter is a static field and is of no interest for the dynamical problems with which we are usually concerned.

of the ψ_n as in (1.7). This connection between the two bases (ψ_n and φ_n) constitutes the unitary transformation.

If C is an operator (matrix) which generates the unitary transformation and φ_1 and φ_2 are any two functions or column matrices (so that $C\varphi_1$ and $C\varphi_2$ are the transformed functions), then, by definition,

$$(C\varphi_1, C\varphi_2) = (\varphi_1, \varphi_2) \tag{1.8'}$$

Using the properties (1.2) of the Hermitian adjoint, this is

$$(\varphi_1, C^+C\varphi_2) = (\varphi_1, \varphi_2)$$

Since φ_1 and φ_2 are any two functions, we conclude that

$$C^+C = 1 \tag{1.8}$$

By a theorem of determinants we can write

$$\det C^+ \det C = \det 1 = 1$$

where det C is the determinant of the matrix C. Since $\det C^+ = (\det C)^*$ it follows that

$$|\det C|^2 = 1$$

But, since $\det C \neq 0$ (that is, it is non-singular), it follows that the matrix of C has an inverse C^{-1} such that $CC^{-1} = C^{-1}C = 1$. Not only can we write the linear equations expressing the new basis in terms of the old, but we can also solve these linear equations and express the old basis in terms of the new. Equation (1.8) is just the relation that defines the inverse[7] of C. That is,

$$C^{-1} = C^+ \tag{1.9}$$

As can be seen from (1.8) or (1.9), the product of any number of unitary matrices (or transformations taken successively) is a unitary matrix (or unitary transformation). Also, if $C = C_1C_2$ and $CC^+ = 1$ and $C_1C_1^+ = 1$, it follows that $1 = CC^+ = C_1C_2C_2^+C_1^+$ or $C_1C_2C_2^+ = C_1$. Since C_1^{-1} exists, we find $C_2C_2^+ = 1$. This is the statement given above with the added result that the inverse of a unitary matrix is also unitary.

The result (1.8) may be written in terms of the matrix elements C_{nm}. These are defined by

$$\psi_n = \sum_m C_{nm}\varphi_m \tag{1.10}$$

where the ψ_n and φ_m are the basis vectors in new and old representations, respectively, and, because of the orthonormality of the φ_m,

$$C_{nm} = (\varphi_m, \psi_n) \tag{1.11}$$

[7] By the definition, if $C = C_1C_2 \cdots C_n$, then $C^{-1} = C_n^{-1} \cdots C_2^{-1}C_1^{-1}$.

Rewriting the operator equations of (1.9), $CC^+ = 1$ and $C^+C = 1$, in matrix notation leads to the unitary properties of the matrix elements C_{mn},

$$\sum_l C_{nl}^* C_{ml} = \delta_{nm} \tag{1.12}$$

$$\sum_l C_{ln}^* C_{lm} = \delta_{nm} \tag{1.13}$$

where we have used the fact that $C_{nm}^+ = C_{mn}^*$. Thus the orthogonality property is obtained whether we sum on the row or the column index. One way of looking at this result is to consider the elements in any row, say the nth, as a "vector" $\mathbf{C}_{n-}$; the lth element of $\mathbf{C}_{n-}$ is C_{nl}. Then (1.12) states that these row vectors are orthonormal: $\mathbf{C}_{n-} \cdot \mathbf{C}_{m-} = \delta_{nm}$. Similarly we may consider the elements of any column as constituting a vector $\mathbf{C}_{-n}$, and find from (1.13) that these column vectors are orthonormal, $\mathbf{C}_{-n} \cdot \mathbf{C}_{-m} = \delta_{nm}$.

Once the transformation from one basis to another has been defined, the next question is to determine how the matrix representation of an operator changes when we change the basis. Thus, for an operator Ω, if we know the elements $(\varphi', \Omega\varphi)$, what are the elements [8] $(\psi', \Omega\psi)$? The matrix element of the operator Ω between two states described by wave functions φ' and φ is

$$(\varphi', \Omega\varphi) \equiv (\varphi' | \Omega | \varphi) \tag{1.14}$$

The transformed functions [9] are $\psi = C\varphi$ and $\psi' = C\varphi'$, and

$$(\varphi', \Omega\varphi) = (C^{-1}\psi', \Omega C^{-1}\psi) = (\psi', (C^{-1})^+ \Omega C^{-1}\psi) = (\psi', \Omega_T \psi)$$

where

$$\Omega_T = (C^{-1})^+ \Omega C^{-1} \tag{1.15}$$

is the transform of the operator Ω. For a unitary transformation $(C^{-1})^+ = C$ and

$$\Omega_T = C\Omega C^{-1} = C\Omega C^+ \tag{1.16}$$

That is, the matrix representation of an operator in the ψ basis is Ω_T when its representation in the φ basis is Ω and these are connected by (1.16). The choice of basis is arbitrary, and the one most convenient

[8] Sometimes we shall use the symbol $(\psi' | \Omega | \psi)$ for the matrix element $(\psi', \Omega\psi)$. This is especially convenient if the quantum numbers or eigenvalues are used for labeling. Thus, $\Omega_{mn} \equiv (\psi_m, \Omega\psi_n) \equiv (\psi_m | \Omega | \psi_n) \equiv (m | \Omega | n)$.

[9] ψ and φ are a column matrices with elements ψ_n and φ_n, respectively; see equation (1.10).

for calculation is often used.[10] Physical results are, of course, independent of this choice, and the preceding discussion expresses this explicitly insofar as the results of measurements are formulated in terms of matrix elements. For example, C may be the unitary transformation for a rotation of the coordinate system chosen to make the calculation of a physical quantity as simple as possible.

3. DIAGONALIZATION OF OPERATORS

The operator Ω is diagonal in the representation of its eigenfunctions

$$\Omega\psi_n = \omega_n\psi_n \tag{1.17}$$

$$(\psi_m, \Omega\psi_n) = \omega_n\delta_{nm}$$

Let Ω_{mn} be its matrix elements in some other basis φ, so that

$$\Omega_{mn} = (\varphi_m, \Omega\varphi_n)$$

$$\Omega\varphi_n = \sum_m \Omega_{mn}\varphi_m \tag{1.18}$$

Starting in the representation of the φ_n, it is an important problem to determine the unitary transformation C from the set φ_n to the set ψ_n which diagonalizes Ω. For simplicity let us assume that Ω is Hermitian with discrete, non-degenerate eigenvalues. This will be the case in many applications of physical interest. From (1.10) we have

$$\psi_n = \sum_m C_{nm}\varphi_m$$

Operate on both sides with Ω; on the left, use (1.17),

$$\Omega\psi_n = \omega_n\psi_n = \omega_n \sum_m C_{nm}\varphi_m$$

and, on the right, use (1.18),

$$\Omega \sum_m C_{nm}\varphi_m = \sum_m C_{nm} \sum_l \Omega_{lm}\varphi_l$$

Equating these two gives

$$\sum_l \sum_m \Omega_{lm}C_{nm}\varphi_l = \omega_n \sum_l C_{nl}\varphi_l$$

[10] This is completely analogous to the arbitrariness involved in the choice of a coordinate system in the description of any physical process. This does not imply that the description of all physical situations can be given without specifying a coordinate system (e.g., angular distributions) but that the choice is only a matter of convenience. In any event, no matter how we choose the coordinate system, the answer to the problem of angular distributions, e.g., relative to a *physically* defined direction (propagation vector of an incident beam), is always the same.