EUGENE P. WIGNER

GROUP THEORY

AND ITS APPLICATION TO THE QUANTUM MECHANICS OF ATOMIC SPECTRA

GROUP THEORY

AND ITS APPLICATION TO THE QUANTUM MECHANICS OF ATOMIC SPECTRA

EUGENE P. WIGNER
Palmer Physical Laboratory, Princeton University
Princeton, New Jersey

TRANSLATED FROM THE GERMAN BY

J. J. GRIFFIN
University of California, Los Alamos Scientific Laboratory
Los Alamos, New Mexico

EXPANDED AND IMPROVED EDITION

1959

ACADEMIC PRESS NEW YORK AND LONDON

Author's Preface

The purpose of this book is to describe the application of group theoretical methods to problems of quantum mechanics with specific reference to atomic spectra. The actual solution of quantum mechanical equations is, in general, so difficult that one obtains by direct calculations only crude approximations to the real solutions. It is gratifying, therefore, that a large part of the relevant results can be deduced by considering the fundamental symmetry operations.

When the original German version was first published, in 1931, there was a great reluctance among physicists toward accepting group theoretical arguments and the group theoretical point of view. It pleases the author that this reluctance has virtually vanished in the meantime and that, in fact, the younger generation does not understand the causes and the basis for this reluctance. Of the older generation it was probably M. von Laue who first recognized the significance of group theory as the natural tool with which to obtain a first orientation in problems of quantum mechanics. Von Laue's encouragement of both publisher and author contributed significantly to bringing this book into existence. I like to recall his question as to which results derived in the present volume I considered most important. My answer was that the explanation of Laporte's rule (the concept of parity) and the quantum theory of the vector addition model appeared to me most significant. Since that time, I have come to agree with his answer that the recognition that almost all rules of spectroscopy follow from the symmetry of the problem is the most remarkable result.

Three new chapters have been added in translation. The second half of Chapter 24 reports on the work of Racah and of his followers. Chapter 24 of the German edition now appears as Chapter 25. Chapter 26 deals with time inversion, a symmetry operation which had not yet been recognized at the time the German edition was written. The contents of the last part of this chapter, as well as that of Chapter 27, have not appeared before in print. While Chapter 27 appears at the end of the book for editorial reasons, the reader may be well advised to glance at it when studying, in Chapters 17 and 24, the relevant concepts. The other chapters represent the translation of Dr. J. J. Griffin, to whom the author is greatly indebted for his ready acceptance of several suggestions and his generally cooperative attitude. He also converted the left-handed coordinate system originally used to a right-handed system and added an Appendix on notations.

The character of the book—its explicitness and its restriction to one subject only, viz. the quantum mechanics of atomic spectra—has not been changed. Its principal results were contained in articles first published in the *Zeitschrift für Physik* in 1926 and early 1927. The initial stimulus for these articles was given by the investigations of Heisenberg and Dirac on the quantum theory of assemblies of identical particles. Weyl delivered lectures in Zürich on related subjects during the academic year 1927–1928. These were later expanded into his well-known book.

When it became known that the German edition was being translated, many additions were suggested. It is regrettable that most of these could not be followed without substantially changing the outlook and also the size of the volume. Author and translator nevertheless are grateful for these suggestions which were very encouraging. The author also wishes to thank his colleagues for many stimulating discussions on the role of group theory in quantum mechanics as well as on more specific subjects. He wishes to record his deep indebtedness to Drs. Bargmann, Michel, Wightman, and, last but not least, J. von Neumann.

E. P. Wigner

Princeton, New Jersey
February, 1959

Translator's Preface

This translation was initiated while the translator was a graduate student at Princeton University. It was motivated by the lack of a good English work on the subject of group theory from the physicist's point of view. Since that time, several books have been published in English which deal with group theory in quantum mechanics. Still, it is perhaps a reasonable hope that this translation will facilitate the introduction of English-speaking physicists to the use of group theory in modern physics.

The book is an interlacing of physics and mathematics. The first three chapters discuss the elements of linear vector theory. The second three deal more specifically with the rudiments of quantum mechanics itself. Chapters 7 through 16 are again mathematical, although much of the material covered should be familiar from an elementary course in quantum theory. Chapters 17 through 23 are specifically concerned with atomic spectra, as is Chapter 25. The remaining chapters are additions to the German text; they discuss topics which have been developed since the original publication of this book: the recoupling (Racah) coefficients, the time inversion operation, and the classical interpretations of the coefficients.

Various readers may wish to utilize the book differently. Those who are interested specifically in the mathematics of group theory might skim over the chapters dealing with quantum physics. Others might choose to de-emphasize the mathematics, touching Chapters 7, 9, 10, 13, and 14 lightly for background and devoting more attention to the subsequent chapters. Students of quantum mechanics and physicists who prefer familiar material interwoven with the less familiar will probably apply a more even distribution of emphasis.

The translator would like to express his gratitude to Professor E. P. Wigner for encouraging and guiding the task, to Drs. Robert Johnston and John McHale who suggested various improvements in the text, and to Mrs. Marjorie Dresback whose secretarial assistance was most valuable.

J. J. GRIFFIN

Los Alamos, New Mexico
February, 1959

ACADEMIC PRESS INC.
111 FIFTH AVENUE
NEW YORK 3, N. Y.

United Kingdom Edition
Published by
ACADEMIC PRESS INC. (LONDON) LTD.
40 PALL MALL, LONDON SW 1

Library of Congress Catalog Card Number 59-10741

PRINTED IN THE UNITED STATES OF AMERICA

Contents

1. Vectors and Matrices

Linear Transformations

An aggregate of n numbers $(\mathfrak{v}_1, \mathfrak{v}_2, \mathfrak{v}_3, \cdots, \mathfrak{v}_n)$ is called an n-dimensional vector, or a vector in n-dimensional space; the numbers themselves are the components of this vector. The coordinates of a point in n-dimensional space can also be interpreted as a vector which connects the origin of the coordinate system with the point considered. Vectors will be denoted by bold face German letters; their components will carry a roman index which specifies the coordinate axis. Thus $\mathfrak{v}_k$ is a vector component (a number), and $\mathfrak{v}$ is a vector, a set of n numbers.

Two vectors are said to be equal if their corresponding components are equal. Thus

$$\mathfrak{v} = \mathfrak{w} \tag{1.1}$$

is equivalent to the n equations

$$\mathfrak{v}_1 = \mathfrak{w}_1; \qquad \mathfrak{v}_2 = \mathfrak{w}_2; \qquad \cdots; \qquad \mathfrak{v}_n = \mathfrak{w}_n .$$

A vector is a null vector if all its components vanish. The product $c\mathfrak{v}$ of a number c with a vector $\mathfrak{v}$ is a vector whose components are c times the components of $\mathfrak{v}$, or $(c\mathfrak{v})_k = c\mathfrak{v}_k$. Addition of vectors is defined by the rule that the components of the sum are equal to the sums of the corresponding components. Formally

$$(\mathfrak{v} + \mathfrak{w})_k = \mathfrak{v}_k + \mathfrak{w}_k . \tag{1.2}$$

In mathemtaical problems it is often advantageous to introduce new variables in place of the original ones. In the simplest case the new variables $x_1', x_2', \cdots, x_n'$ are linear functions of the old ones, $x_1, x_2, \cdots, x_n$. That is

$$\begin{aligned} x_1' &= \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n \\ x_2' &= \alpha_{21}x_1 + \cdots + \alpha_{2n}x_n \\ &\cdot \\ &\cdot \\ &\cdot \\ x_n' &= \alpha_{n1}x_1 + \cdots + \alpha_{nn}x_n \end{aligned} \tag{1.3}$$

or

$$x_i' = \sum_{k=1}^{n} \alpha_{ik}x_k . \tag{1.3a}$$

The introduction of new variables in this way is called *linear transformation*.

The transformation is completely determined by the coefficients $\alpha_{11}, \cdots, \alpha_{nn}$, and the aggregate of these n^2 numbers arranged in a square array is called the *matrix* of the linear transformation (1.3):

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix} \tag{1.4}$$

We shall write such a matrix more concisely as (α_{ik}) or simply $\boldsymbol{\alpha}$.

For Eq. (1.3) actually to represent an introduction of new variables, it is necessary not only that the x' be expressible in terms of the x, but also that the x can be expressed in terms of the x'. That is, if we view the x_i as unknowns in Eq. (1.3), a unique solution to these equations must exist giving the x in terms of the x'. The necessary and sufficient condition for this is that the determinant formed from the coefficients α_{ik} be nonzero:

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{vmatrix} \neq 0. \tag{1.4a}$$

Transformations whose matrices have nonvanishing determinants are referred to as *proper* transformations, but an array of coefficients like (1.4) is always called a matrix, whether or not it induces a proper transformation. Boldface letters are used to represent matrices; matrix coefficients are indicated by affixing indices specifying the corresponding axes. Thus $\boldsymbol{\alpha}$ is a matrix, an array of n^2 numbers; α_{jk} is a matrix element (a number).

Two matrices are equal if all their corresponding coefficients are equal. Thus

$$\boldsymbol{\alpha} = \boldsymbol{\beta} \tag{1.5}$$

is equivalent to the n^2 equations

$$\alpha_{jk} = \beta_{jk} \qquad (j, k = 1, 2, \cdots, n).$$

Another interpretation can be placed on the equation

$$x_i' = \sum_{k=1}^{n} \alpha_{ik} x_k \tag{1.3a}$$

by considering the x_j', not as components of the original vector in a new coordinate system, but as the components of a *new vector in the original* *coordinate system.* We then say that the matrix α transforms the vector $\mathfrak{x}$ into the vector $\mathfrak{x}'$, or that α applied to $\mathfrak{x}$ gives $\mathfrak{x}'$

$$\mathfrak{x}' = \alpha\mathfrak{x}. \tag{1.3b}$$

This equation is completely equivalent to (1.3a).

An n-dimensional matrix is a *linear operator* on n-dimensional vectors. It is an *operator* because it transforms one vector into another vector; it is *linear* since for arbitrary numbers a and b, and arbitrary vectors $\mathfrak{r}$ and $\mathfrak{v}$, the relation

$$\alpha(a\mathfrak{r} + b\mathfrak{v}) = a\alpha\mathfrak{r} + b\alpha\mathfrak{v} \tag{1.6}$$

is true. To prove (1.6) one need only write out the left and right sides explicitly. The kth component of $a\mathfrak{r} + b\mathfrak{v}$ is $a\mathfrak{r}_k + b\mathfrak{v}_k$, so that the ith component of the vector on the left is:

$$\sum_{k=1}^{n} \alpha_{ik}(a\mathfrak{r}_k + b\mathfrak{v}_k).$$

But this is identical with the ith component of the vector on the right side of (1.6)

$$a\sum_{k=1}^{n} \alpha_{ik}\mathfrak{r}_k + b\sum_{k=1}^{n} \alpha_{ik}\mathfrak{v}_k .$$

This establishes the linearity of matrix operators.

An n-dimensional matrix is the *most general* linear operator in n-dimensional vector space. That is, every linear operator in this space is equivalent to a matrix. To prove this, consider the arbitrary linear operator $\mathbf{O}$ which transforms the vector $\mathfrak{e}_1 = (1, 0, 0, \cdots, 0)$ into the vector $\mathfrak{r}_{\cdot 1}$, the vector $\mathfrak{e}_2 = (0, 1, 0, \cdots, 0)$ into the vector $\mathfrak{r}_{\cdot 2}$, and finally, the vector $\mathfrak{e}_n = (0, 0, 0, \cdots, 1)$ into $\mathfrak{r}_{\cdot n}$, where the components of the vector $\mathfrak{r}_{\cdot k}$ are $\mathfrak{r}_{1k}, \mathfrak{r}_{2k}, \cdots, \mathfrak{r}_{nk}$. Now the matrix $(\mathfrak{r}_{ik})$ transforms each of the vectors $\mathfrak{e}_1, \mathfrak{e}_2, \cdots, \mathfrak{e}_n$ into the same vectors, $\mathfrak{r}_{\cdot 1}, \mathfrak{r}_{\cdot 2}, \cdots, \mathfrak{r}_{\cdot n}$ as does the operator $\mathbf{O}$. Moreover, any n-dimensional vector $\mathfrak{a}$ is a linear combination of the vectors $\mathfrak{e}_1, \mathfrak{e}_2, \cdots, \mathfrak{e}_n$. Thus, both $\mathbf{O}$ and $(\mathfrak{r}_{ik})$ (since they are linear) transform any arbitrary vector $\mathfrak{a}$ into the same vector $a_1\mathfrak{r}_{\cdot 1} + \cdots + a_n\mathfrak{r}_{\cdot n}$. The matrix $(\mathfrak{r}_{ik})$ is therefore equivalent to the operator $\mathbf{O}$.

The most important property of linear transformations is that two of them, applied successively, can be combined into a single linear transformation. Suppose, for example, we introduce the variables x' in place of the original x via the linear transformation (1.3), and subsequently introduce variables x'' via *a second linear transformation,*

$$\begin{aligned} x_1'' &= \beta_{11}x_1' + \beta_{12}x_2' + \cdots + \beta_{1n}x_n' \\ &\cdot \qquad \cdot \qquad\qquad\qquad \cdot \\ &\cdot \qquad \cdot \qquad\qquad\qquad \cdot \\ &\cdot \qquad \cdot \qquad\qquad\qquad \cdot \\ x_n'' &= \beta_{n1}x_1' + \beta_{n2}x_2' + \cdots + \beta_{nn}x_n'. \end{aligned} \tag{1.7}$$

Both processes can be *combined into a single one*, so that the x'' are introduced directly in place of the x by one linear transformation. Substituting (1.3) into (1.7), one finds

$$\begin{aligned} x''_1 &= \beta_{11}(\alpha_{11}x_1 + \cdots + \alpha_{1n}x_n) + \cdots + \beta_{1n}(\alpha_{n1}x_1 + \cdots + \alpha_{nn}x_n) \\ x''_2 &= \beta_{21}(\alpha_{11}x_1 + \cdots + \alpha_{1n}x_n) + \cdots + \beta_{2n}(\alpha_{n1}x_1 + \cdots + \alpha_{nn}x_n) \\ &\cdot \\ &\cdot \\ &\cdot \\ x''_n &= \beta_{n1}(\alpha_{11}x_1 + \cdots + \alpha_{1n}x_n) + \cdots + \beta_{nn}(\alpha_{n1}x_1 + \cdots + \alpha_{nn}x_n). \end{aligned} \tag{1.8}$$

Thus, the x'' are linear functions of the x. We can write (1.8) more concisely by condensing (1.3) and (1.7)

$$x'_j = \sum_{k=1}^{n} \alpha_{jk}x_k \qquad (j = 1, 2, \cdots, n) \tag{1.3c}$$

$$x''_i = \sum_{j}^{n} \beta_{ij}x'_j \qquad (i = 1, 2, \cdots, n). \tag{1.7a}$$

Then (1.8) becomes

$$x''_i = \sum_{j=1}^{n} \sum_{k=1}^{n} \beta_{ij}\alpha_{jk}x_k. \tag{1.8a}$$

Furthermore, by defining γ through

$$\gamma_{ik} = \sum_{j=1}^{n} \beta_{ij}\alpha_{jk} \tag{1.9}$$

one obtains simply

$$x''_i = \sum_{k=1}^{n} \gamma_{ik}x_k. \tag{1.8b}$$

This demonstrates that the combination of two linear transformations (1.7) and (1.3), with matrices (β_{ik}) and (α_{ik}) is a single linear transformation which has the matrix (γ_{ik}).

The matrix (γ_{ik}), defined in terms of the matrices (α_{ik}) and (β_{ik}) according to Eq. (1.9), is called the *product* of the matrices (β_{ik}) and (α_{ik}). Since (α_{ik}) transforms the vector $\mathfrak{r}$ into $\mathfrak{r}' = \alpha\mathfrak{r}$, and (β_{ik}) transforms the vector $\mathfrak{r}'$ into $\mathfrak{r}'' = \beta\mathfrak{r}'$, the product matrix (γ_{ik}) by its definition, transforms $\mathfrak{r}$ directly into $\mathfrak{r}'' = \gamma\mathfrak{r}$. This method of combining transformations is called "matrix multiplication," and exhibits a number of simple properties, which we now enumerate as theorems.

First of all we observe that the formal rule for matrix multiplication is the same as the rule for the multiplication of determinants.

1. *The determinant of a product of two matrices is equal to the product of the determinants of the two factors.*

In the multiplication of matrices, it is not necessarily true that

$$\alpha\beta = \beta\alpha. \tag{1.E.1}$$

For example, consider the two matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

This establishes a second property of matrix multiplication.

2. *The product of two matrices depends in general upon the order of the factors.*

In the very special situation when Eq. (1.E.1) is true, the matrices α and β are said to *commute*.

In contrast to the commutative law,

3. *The associative law of multiplication is valid in matrix multiplication.*

That is,

$$\gamma(\beta\alpha) = (\gamma\beta)\alpha. \tag{1.10}$$

Thus, it makes no difference whether one multiplies γ with the product of β and α, or the product of γ and β with α. To prove this, denote the i–kth element of the matrix on the left side of (1.10) by ϵ_{ik}. Then

$$\epsilon_{ik} = \sum_{j=1}^{n} \gamma_{ij}(\beta\alpha)_{jk} = \sum_{j=1}^{n} \sum_{l=1}^{n} \gamma_{ij}\beta_{jl}\alpha_{lk}. \tag{1.10a}$$

The i–kth element on the right side of (1.10) is

$$\epsilon'_{ik} = \sum_{l=1}^{n} (\gamma\beta)_{il}\alpha_{lk} = \sum_{l=1}^{n} \sum_{j=1}^{n} \gamma_{ij}\beta_{jl}\alpha_{lk}. \tag{1.10b}$$

Then $\epsilon_{ik} = \epsilon'_{ik}$, and (1.10) is established. One can therefore write simply $\gamma\beta\alpha$ for both sides of (1.10).

The validity of the associative law is immediately obvious if the matrices are considered as linear operators. Let α transform the vector $\mathfrak{r}$ into $\mathfrak{r}' = \alpha\mathfrak{r}$, β the vector $\mathfrak{r}'$ into $\mathfrak{r}'' = \beta\mathfrak{r}'$, and γ the vector $\mathfrak{r}''$ into $\mathfrak{r}''' = \gamma\mathfrak{r}''$. Then the combination of two matrices into a single one by matrix multiplication signifies simply the combination of two operations. The product $\beta\alpha$ transforms $\mathfrak{r}$ directly into $\mathfrak{r}''$, and $\gamma\beta$ transforms $\mathfrak{r}'$ directly into $\mathfrak{r}'''$. Thus both $(\gamma\beta)\alpha$ and $\gamma(\beta\alpha)$ transform $\mathfrak{r}$ into $\mathfrak{r}'''$, and the two operations are equivalent.

4. *The unit matrix*

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \tag{1.11}$$

plays a special role in matrix multiplication, just as the number 1 does in ordinary multiplication. For every matrix $\boldsymbol{\alpha}$,

$$\boldsymbol{\alpha} \cdot \mathbf{1} = \mathbf{1} \cdot \boldsymbol{\alpha}.$$

That is, $\mathbf{1}$ commutes with all matrices, and its product with any matrix is just that matrix again. The elements of the unit matrix are denoted by the symbol δ_{ik}, so that

$$\delta_{ik} = 0 \qquad (i \neq k)$$

$$\delta_{ik} = 1 \qquad (i = k). \tag{1.12}$$

The δ_{ik} defined in this way is called the Kronecker delta-symbol. The matrix $(\delta_{ik}) = \mathbf{1}$ induces the *identity* transformation, which leaves the variables unchanged.

If for a given matrix $\boldsymbol{\alpha}$, there exists a matrix $\boldsymbol{\beta}$ such that

$$\boldsymbol{\beta\alpha} = \mathbf{1}, \tag{1.13}$$

then $\boldsymbol{\beta}$ is called the *inverse*, or *reciprocal*, of the matrix $\boldsymbol{\alpha}$. Equation (1.13) states that a transformation via the matrix $\boldsymbol{\beta}$ exists which combines with $\boldsymbol{\alpha}$ to give the identity transformation. If the determinant of $\boldsymbol{\alpha}$ is not equal to zero ($|\alpha_{ik}| \neq 0$), then an inverse transformation always exists (as has been mentioned on page 2). To prove this we write out the n^2 equations (1.13) more explicitly

$$\sum_{j=1}^{n} \beta_{ij}\alpha_{jk} = \delta_{ik} \qquad (i, k = 1, 2, \cdots, n). \tag{1.14}$$

Consider now the n equations in which i has one value, say l. These are n linear equations for n unknowns $\beta_{l1}, \beta_{l2}, \cdots, \beta_{ln}$. They have, therefore, one and only one solution, provided the determinant $|\alpha_{jk}|$ does not vanish. The same holds for the other $n - 1$ systems of equations. This establishes the fifth property we wish to mention.

5. *If the determinant* $|\alpha_{jk}| \neq 0$, *there exists one and only one matrix* $\boldsymbol{\beta}$ *such that* $\boldsymbol{\beta\alpha} = \mathbf{1}$.

Moreover, the determinant $|\beta_{jk}|$ is the reciprocal of $|\alpha_{jk}|$, since, according to Theorem 1,

$$|\beta_{jk}| \cdot |\alpha_{jk}| = |\delta_{jk}| = 1. \tag{1.15}$$

From this it follows that α *has no inverse if* $|\alpha_{ik}| = 0$, and that β, the inverse of α, must also have an inverse.

We now show that if (1.13) is true, then

$$\alpha\beta = 1 \tag{1.16}$$

is true as well. That is, if β is the inverse of α, then α is also the inverse of β. This can be seen most simply by multiplying (1.13) from the right with β,

$$\beta\alpha\beta = \beta, \tag{1.17}$$

and this from the left with the inverse of β, which we call γ. Then

$$\gamma\beta\alpha\beta = \gamma\beta$$

and since, by hypothesis $\gamma\beta = 1$, this is identical with (1.16). Conversely, (1.13) follows easily from (1.16). This proves Theorem 6 (the inverse of α is denoted by α^{-1}).

6. *If α^{-1} is the inverse of α, then α is also the inverse of α^{-1}.*

It is clear that *inverse matrices commute* with one another.

Rule: The inverse of a product $\alpha\beta\gamma\delta$ is obtained by multiplying the inverses of the individual factors in reverse order $(\delta^{-1}\gamma^{-1}\beta^{-1}\alpha^{-1})$. That is

$$(\delta^{-1}\gamma^{-1}\beta^{-1}\alpha^{-1}) \cdot (\alpha\beta\gamma\delta) = 1.$$

Another important matrix is

7. *The null matrix, every element of which is zero.*

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ . & . & . & \cdots & . \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{1.18}$$

Obviously one has

$$\alpha \cdot \mathbf{0} = \mathbf{0} \cdot \alpha = \mathbf{0}$$

for any matrix α.

The null matrix plays an important role in another combination process for matrices, namely, addition. The sum γ of two matrices α and β is the matrix whose elements are

$$\gamma_{ik} = \alpha_{ik} + \beta_{ik}. \tag{1.19}$$

The n^2 equations (1.19) are equivalent to the equation

$$\gamma = \alpha + \beta \quad \text{or} \quad \gamma - \alpha - \beta = \mathbf{0}.$$

Addition of matrices is clearly commutative.

$$\boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\beta} + \boldsymbol{\alpha}. \tag{1.20}$$

Moreover, multiplication by sums is distributive.

$$\boldsymbol{\gamma}(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \boldsymbol{\gamma\alpha} + \boldsymbol{\gamma\beta}$$
$$(\boldsymbol{\alpha} + \boldsymbol{\beta})\boldsymbol{\gamma} = \boldsymbol{\alpha\gamma} + \boldsymbol{\beta\gamma}.$$

Furthermore, the product of a matrix $\boldsymbol{\alpha}$ and a number a is defined to be that matrix $\boldsymbol{\gamma}$ each element of which is a times the corresponding elements of $\boldsymbol{\alpha}$.

$$\gamma_{ik} = a\alpha_{ik}. \tag{1.21}$$

The formulas

$$(ab)\boldsymbol{\alpha} = a(b\boldsymbol{\alpha}); \qquad \boldsymbol{\alpha}a\boldsymbol{\beta} = a\boldsymbol{\alpha\beta}; \qquad a(\boldsymbol{\alpha} + \boldsymbol{\beta}) = a\boldsymbol{\alpha} + a\boldsymbol{\beta}$$

then follow directly.

Since integral powers of a matrix $\boldsymbol{\alpha}$ can easily be defined by successive multiplication

$$\boldsymbol{\alpha}^2 = \boldsymbol{\alpha} \cdot \boldsymbol{\alpha}; \qquad \boldsymbol{\alpha}^3 = \boldsymbol{\alpha} \cdot \boldsymbol{\alpha} \cdot \boldsymbol{\alpha}; \ldots$$
$$\boldsymbol{\alpha}^{-2} = \boldsymbol{\alpha}^{-1} \cdot \boldsymbol{\alpha}^{-1}; \qquad \boldsymbol{\alpha}^{-3} = \boldsymbol{\alpha}^{-1} \cdot \boldsymbol{\alpha}^{-1} \cdot \boldsymbol{\alpha}^{-1}; \ldots \tag{1.22}$$

polynomials with positive and negative integral exponents can also be defined

$$\cdots + a_{-n}\boldsymbol{\alpha}^{-n} + \cdots + a_{-1}\boldsymbol{\alpha}^{-1} + a_0\mathbf{1} + a_1\boldsymbol{\alpha} + \cdots + a_n\boldsymbol{\alpha}^n + \cdots. \tag{1.23}$$

The coefficients a in the above expression are not matrices, but numbers. *A function of* $\boldsymbol{\alpha}$ *like* (1.23) *commutes with any other function of* $\boldsymbol{\alpha}$ (*and, in particular, with* $\boldsymbol{\alpha}$ *itself*).

Still another important type of matrix which appears frequently is the diagonal matrix.

8. *A diagonal matrix is a matrix the elements of which are all zero except for those on the main diagonal.*

$$\mathbf{D} = \begin{pmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot\ \cdot \\ 0 & 0 & \cdots & D_n \end{pmatrix}. \tag{1.24}$$

The general element of this diagonal matrix can be written

$$\mathbf{D}_{ik} = D_i\delta_{ik}. \tag{1.25}$$

All diagonal matrices commute, and the product of two diagonal matrices is again diagonal. This can be seen directly from the definition of the product.

$$(\mathbf{DD}')_{ik} = \sum_j \mathbf{D}_{ij}\mathbf{D}'_{jk} = \sum_j D_i\delta_{ij}D'_j\delta_{jk} = D_iD_i{}'\delta_{ik}. \tag{1.26}$$