QUANTUM THEORY PROJECT
FOR RESEARCH IN ATOMIC, MOLECULAR AND SOLID STATE
CHEMISTRY AND PHYSICS
UNIVERSITY OF FLORIDA, GAINESVILLE, FLORIDA

LINEAR ALGEBRA AND THE FUNDAMENTS OF QUANTUM THEORY

by

Per-Olov Löwdin

PREPRINT NO. 65
July 1, 1964

This work was supported in part by the
National Aeronautics and Space Administration
under Research Grant No. G-512 with the
University of Florida

QUANTUM THEORY PROJECT
FOR RESEARCH IN ATOMIC, MOLECULAR AND SOLID STATE
CHEMISTRY AND PHYSICS
UNIVERSITY OF FLORIDA, GAINESVILLE, FLORIDA

LINEAR ALGEBRA AND THE FUNDAMENTS OF QUANTUM THEORY*

by

Per-Olov Löwdin

PREPRINT NO. 65
July 1, 1964

* Since the author is also connected with the University of Uppsala, this report has been issued as TN No. 125 from the Quantum Chemistry Group.

This work was supported in part by the
National Aeronautics and Space Administration
under Research Grant No. G-512 with the
University of Florida

CONTENTS

ABSTRACT

17083

It is investigated how far one can go in the formulation of the fundaments of quantum theory by using the axioms of linear algebra alone, i.e. without the help of the concept of the scalar product. In treating linear spaces of finite order, one can formulate the eigenvalue problem for a linear operator and reach the concepts of matrix representation, eigenexpansions, and transformation to diagonal and classical canonical form. One can further define the concepts of projection operators, the resolution of the identity, and the spectral resolution of an operator.

In treating infinite spaces, the interest is confined to operators having all their eigenvalues situated in a finite number of points in the complex plane, each of which may be infinitely degenerate. Assuming that the operator under consideration satisfies a reduced Cayley-Hamilton equation of finite order, one shows that there exists a set of projection operators forming a resolution of the identity by means of which one can carry out a unique "component analysis" of an arbitrary element of the space. Even the spectral resolution of the operator exists. These theorems are in quantum theory of particular importance in treating constants of motion.

In the last section, the scalar product is introduced, and the connection with the conventional approach is studied.

Author

1. INTRODUCTION

When modern quantum theory was introduced in 1925, there were three independent and competing forms, namely the wave mechanics developed by Schrödinger[1], the matrix mechanics introduced by Heisenberg, Born, and Jordan, and the q-number theory developed by Dirac[3]. The equivalence between the three approaches was shown by Schrödinger[4]. In these connections, the concepts of linear operators, linear spaces, and vector spaces play a fundamental role. The physical interpretations of quantum theory are based on the concept of the "expectation value" which is essentially the scalar product of two vectors. The most thorough discussions of the foundations of quantum theory given by von Neumann[5] and by Dirac[6] are using the concepts of the theory of Hilbert space in which the scalar product plays a basic role.

Of fundamental importance in quantum theory are the "constants of motion", i.e. physical quantities which are associated with linear operators Λ commuting with the Hamiltonian H, so that $H\Lambda = \Lambda H$. They have eigenvalue problems of the type

$$\Lambda \Phi_k = \lambda_k \Phi_k ,$$

and the eigenvalues λ_k are used to classify the stationary states and also certain properties of time-dependent phenomena. In atomic theory, typical constants of motion are represented by the total spin $\vec{S}$, the orbital angular momentum $\vec{L}$, and the total angular momentum $\vec{J} = \vec{L} + \vec{S}$ in various coupling schemes. In a study of certain classes of constants of motion, the author has developed a technique based on the use of product-type projection operators[7], which has turned out to be rather useful in practical applications[8]. In a survey of this method[9] for normal constants of motion Λ, satisfying the relation $\Lambda^\dagger\Lambda = \Lambda\Lambda^\dagger$, it has been shown that eigenfunctions Φ_k and Φ_ℓ associated with different eigenvalues λ_k and λ_ℓ, respectively, are not only orthogonal but also non-interacting with respect to H, so that

$$\langle \Phi_k | \Phi_\ell \rangle = 0 , \qquad \langle \Phi_k | \mathcal{H} | \Phi_\ell \rangle = 0 , \qquad \lambda_k \neq \lambda_\ell$$

Some fundamental theorems as "the resolution of the identity" and the "spectral resolution of Λ" were also demonstrated in an elementary way, but the entire formalism was based on the use of scalar products $\langle | \rangle$.

The purpose of this paper is to generalize these results and to show that essentially the same type of projection-operator formalism may be derived in a theory of linear space alone, i.e. without the use of the concept of scalar products. The resolution of the identity corresponds now to a "component analysis" which, under certain conditions, is valid also for an infinite linear space. It is interesting to see how many fundamental quantum-mechanical theorems may be found and illustrated in this way. Of course, quantum theory will not be complete without the concept of the scalar product and the convergence properties of the Hilbert space, but our approach shows how far one can actually proceed without these important ingredients.

2. LINEAR SPACES AND LINEAR OPERATORS

Linear Spaces

Definitions. - Let us consider a set of elements A, B, C, D, ... which may be subject to two operations called "addition" and "multiplication by a complex constant α" leading to new elements of the form $A + B$ and $\alpha \cdot A$, respectively. The operations are assumed to satisfy the following rules:

$A + B = B + A$,	Commutative law of addition.
$(A+B) + C = A + (B+C)$,	Associative law of addition.
$(\alpha+\beta) A = \alpha A + \beta A$,	First distributive law of multiplication.
$\alpha(A+B) = \alpha A + \alpha B$,	Second distributive law of multiplication.
$(\alpha\beta) A = \alpha(\beta A)$,	Associative law of multiplication.

(1)

A set of this type is called a linear set. If further the set is closed under the two operations, the set is said to form a linear space. This implies that application of the two operations to any elements of the set leads again to an element of the same set. Some simple examples of linear spaces are provided by the following list:

Set of all vectors of a given dimension.
Set of all polynomials of degree equal to or less than n.
Set of all continuous functions.
Set of all integrable functions.

Some elementary rules. - For $\alpha = 0$, one has particularly $0 \cdot A = \bar{0}$, where $\bar{0}$ is called the "zero-element" of the set, which is an independent concept clearly distinct from the complex number 0. For $\alpha = 1$, one has further the rule $1 \cdot A = A$, which gives an important property of the multiplication. By using these rules, one can now prove the elementary theorem

$$A + \overline{0} = A \quad , \tag{2}$$

since one has $A + \overline{0} = 1 \cdot A + 0 \cdot A = (1 + 0) \cdot A = 1 \cdot A = A$ according to (1). Another simple theorem says that, if $\alpha A = \beta B$ and $\alpha \neq 0$, then $A = (\beta/\alpha) B$. The proof follows from the fact that $A = 1 \cdot A =$ $= (\alpha^{-1} \cdot \alpha) A = \alpha^{-1} (\alpha A) = \alpha^{-1} (\beta B) = (\alpha^{-1} \beta) B = (\beta / \alpha) B$, and it is a good illustration how the rules in (1) are applied one by one. As a corollary follows that, if $\alpha A = \overline{0}$ for $\alpha \neq 0$, then $A = \overline{0}$. The operation of "subtraction" is defined by the rule

$$A - B = A + (-1) B \tag{3}$$

One gets immediately the theorem $A - A = \overline{0}$, since one has $A - A = 1 \cdot A + (-1) \cdot A = \{ 1 + (-1) \} A = 0 \cdot A = \overline{0}$. In the same way, one can then proceed to derive a series of elementary arithmetical rules for the linear set of a well-known character.

Linear independence. - Let us now introduce a fundamental concept in the theory of linear spaces by the following definition:

> A finite subset of non-zero elements $A_1, A_2, \ldots A_N$ is said to be linearly independent, if and only if the relation
>
> $$A_1 \alpha_1 + A_2 \alpha_2 + \cdots + A_N \alpha_N = \overline{0} \tag{4}$$
>
> necessarily implies that $\alpha_1 = \alpha_2 = \cdots = \alpha_N = 0$.

This concept provides a tool for going from an arithmetical statement about elements of the linear space to a corresponding statement about complex numbers, and it will, in the following, frequently be used for this purpose. Any subset of elements which is not linearly independent is said to be linearly dependent, and there exists then a linear relation (4) between the elements in which at least two of the coefficients α_k are different from zero.

The next definition deals with the concept of a "basis" of a linear space:

> A set of linearly independent elements $X_1, X_2, \ldots X_n$ is said to form a <u>basis</u> of the linear space, if and only if the subset $A, X_1, X_2, \ldots X_n$ is linearly dependent for every non-zero element A of the linear space; the number n is called the order of the basis.

This definition leads directly to the following "expansion theorem":

> If a linear space has a basis, any element A of the space can be written as a sum
>
> $$A = X_1 a_1 + X_2 a_2 + \cdots + X_n a_n \quad (5)$$

The theorem is trivially true for the zero-element which corresponds to the coefficients $a_1 = a_2 = \ldots = a_n = 0$. For $A \neq \bar{0}$, we will consider the relation

$$A\alpha + X_1\alpha_1 + X_2\alpha_2 + \cdots + X_n\alpha_n = \bar{0}, \quad (6)$$

where now at least two coefficients α_k are different from zero. One has $\alpha \neq 0$, since otherwise the elements of the basis would be linearly dependent, and, multiplying by α^{-1} and putting $a_k = -\alpha^{-1}\alpha_k$, one obtains expansion (5). Next one has the "uniqueness theorem":

> The coefficients a_k in the expansion of a given element A in terms of a basis $X_1, X_2, \ldots X_n$ are unique.

To prove the theorem, let us assume that there are two different expansions of an element A in terms of a given basis, so that

$$A = X_1 a_1 + X_2 a_2 + \cdots + X_n a_n,$$
$$A = X_1 a_1' + X_2 a_2' + \cdots + X_n a_n'.$$

By subtraction, we obtain

$$X_1(a_1 - a_1') + X_2(a_2 - a_2') + \cdots + X_n(a_n - a_n') = \bar{0}, \quad (7)$$

and, since the subset $X_1, X_2, \ldots X_n$ was assumed to be linearly independent, this gives $(a_k - a'_k) = 0$ or

$$a'_k = a_k \tag{8}$$

for all k, which proves the uniqueness theorem.

In the following, it will often be convenient to use "matrix notations" in which bold-face symbols will denote rectangular or quadratic arrangements of elements or complex numbers, so that $\mathbf{K} = \{K_{kl}\}$. A rectangular matrix which consists of a singel row or column will be called a row-vector or column-vector, respectively. A matrix product will further be defined as a new matrix in which the elements are the "inner products" of the rows of the first matrix times the columns of the second matrix:

$$(\mathbf{K} \cdot \mathbf{L})_{kl} = \sum_{\alpha} K_{k\alpha} L_{\alpha l} . \tag{9}$$

The concept is, of course, subject to the compatibility condition that the first matrix should have as many columns as the second has rows. We will further let $\widetilde{\mathbf{K}}$ denote the "transpose" of the matrix $\mathbf{K}$, i.e. the matrix having the rows and columns interchanged, so that $\widetilde{K}_{k\ell} = K_{\ell k}$.

Introducing the row vector $\mathbf{X} = (X_1, X_2, \ldots X_n)$ of the basic elements and the column vector $\mathbf{a}$ of the coefficients a_k:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} , \tag{10}$$

one can hence, instead of (5), simply use the short-hand notation

$$A = \mathbf{X}\,\mathbf{a} . \tag{11}$$

Since the coefficients a_k according to (8) are uniquely defined, we will further introduce the notation

$$a_k = \{X_k, \mathbf{X} \mid A\} , \tag{12}$$

where the right-hand member implies that we consider the expansion of the element A in terms of the basis $\mathbf{X}$ and selects the coefficient associated with X_k. Of course, the symbol does not contain any recipe for the evaluation of the coefficient a_k, and we will return to this question later.

Transformations of basis. - Let a linear space have a basis $\mathbf{X} = (X_1, X_2, \ldots X_n)$, and let us consider an arbitrary subset $\mathbf{Y} = (Y_1, Y_2, \ldots Y_n)$ of linearly independent elements. We will now show that also the set $\mathbf{Y}$ may be used as a basis. For this purpose, we will expand each one of the elements Y_k in terms of $\mathbf{X}$ according to (5), so that

$$Y_k = \sum_l X_l \alpha_{lk} \tag{13}$$

The coefficients in these expansions for $k = 1, 2, \ldots n$ form together a matrix $\boldsymbol{\alpha} = \{\alpha_{lk}\}$, so that one can condense the equations (13) into the form

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\alpha} \tag{14}$$

In the following, it is often convenient to use the theory of determinants. Let $D = \det \boldsymbol{\alpha} = |\alpha_{kl}|$ be the determinant of the matrix $\boldsymbol{\alpha}$, and let Γ_{kl} be the cofactor of the element α_{kl} . The expansion theorem for determinants gives then

$$\begin{aligned} \sum_k \alpha_{lk} \Gamma_{lk} &= D, \\ \sum_k \alpha_{lk} \Gamma_{pk} &= 0 \qquad p \neq l \end{aligned} \tag{15}$$

If $\boldsymbol{\Gamma}$ is the matrix of the elements Γ_{kl} , one can hence write the two relations (15) in the condensed form

$$\boldsymbol{\alpha} \cdot \widetilde{\boldsymbol{\Gamma}} = D \cdot \mathbf{1} \tag{16}$$

Multiplying (14) to the right by $\tilde{\Gamma}$, one obtains $\mathbf{Y}\tilde{\Gamma} = \mathbf{X}\alpha\tilde{\Gamma} = \mathbf{X}\cdot D$. It is clear that $D \neq 0$, since otherwise the set $\mathbf{Y}$ would be linearly dependent, and we note that $D \neq 0$ is a necessary and sufficient condition for the linear independence of the subset $\mathbf{Y}$. Introducing the new matrix $\beta = D^{-1}\tilde{\Gamma} = \alpha^{-1}$, we obtain

$$\mathbf{X} = \mathbf{Y}\beta \tag{17}$$

Substitution of this expression into (11) gives $A = \mathbf{X}\mathbf{a} = \mathbf{Y}\beta\mathbf{a}$, which indicates that the subset $\mathbf{Y}$ may be used as a basis. One gets hence the following transformation formula

$$A = \mathbf{Y}\mathbf{b}, \qquad \mathbf{b} = \beta\mathbf{a} \tag{18}$$

under a change of the basis.

It is now clear that every linearly independent subset of order n may be used as a basis. This shows also that it is impossible to find a basis of another order m, say the linearly independent subset $Z_1, Z_2, \ldots Z_m$, where $m > n$. The elements $Z_1, Z_2, \ldots Z_n$ would again form a basis, in which the remaining elements $Z_{n+1}, Z_{n+2}, \ldots Z_m$ could be expressed, and the subset $Z_1, Z_2, \ldots Z_m$ could then not be linearly independent. In the same way, one proves that the assumption $n > m$ leads to a contradiction, and one has consequently $m = n$. The number of a basis is hence unique and, since it is characteristic for the linear space concerned, it is called the <u>order of the space.</u>

The set of all three-dimensional vectors has, of course, the order three, whereas the space of all polynomials in the variable x of degree less than or equal to n has the order $(n+1)$. As a basis for a description of this space one may choose e.g. the powers $1, x, x^2, \ldots x^n$.

In the first part of our treatment, we will confine our interest to linear spaces of a <u>finite order</u> n, whereas later certain theorems will be generalized also to spaces of an infinite order.

Linear manifolds, - In our study of the linear spaces, it is often convenient to use the concept of the linear manifold introduced by the following definition:

> If $\mathbf{f} = (f_1, f_2, \ldots f_k)$ is a linearly independent set of elements of the linear space, then the collection of all elements $f_1\alpha_1 + f_2\alpha_2 + \cdots f_k\alpha_k$ for arbitrary values of the complex parameters $\alpha_1, \alpha_2, \ldots \alpha_k$ forms a subspace of order k, which is called the linear manifold spanned by the elements $f_1, f_2, \ldots f_k$.

One has always $k \leq n$, but it is usually convenient to reserve the terms given above to the case $k < n$.

From the geometrical point of view, one could speak of a single element f_1 as a "point" in the linear space, whereas the linear manifolds $f_1\alpha_1$ and $f_1\alpha_1 + f_2\alpha_2$ form a "line" and a "plane", respectively.

Linear Operators

An operator T is a rule by means of which one maps the elements A of a linear space onto the elements B of another linear space, so that $B = TA$. The operator concept is apparently a generalization of the idea of a "function" $y = f(x)$, by means of which an independent variable x is mapped onto a dependent variable y. There is one particularly important class of operators characterized by the following definition:

> An operator T is said to be a linear operator, if it satisfies the following two conditions:

$$\begin{aligned} T(A_1 + A_2) &= TA_1 + TA_2, \\ T(\alpha A) &= \alpha\, TA. \end{aligned} \tag{19}$$

The elements A for which the operator T is defined are said to form the domain of T, whereas the elements $B = TA$ are said to form the range of T.

In this section, we will consider only operators mapping a linear space onto itself or onto a subspace of itself, but later we will also study more general mappings.

There are two elementary operators of particular interest, namely the identity operator I and the zero-operator 0_{op} defined by the relations

$$IA = A, \tag{20}$$

$$0_{op}A = \bar{0}, \tag{21}$$

for every element A in the linear space. The concept of the "zero-operator" is, of course, different from the concept of the "zero-element", and we note that the zero-operator is identical with the multiplication by the complex number 0. Important examples of linear operators are given by the differentiation d/dx, the integration $\int^x dx$, and the multiplication by a complex constant α. It is clear that the domains of the first two operations may not coincide with the entire linear space under consideration.

Let us now consider two linear operators F and G. Their sum and product are defined by the relations:

$$(F+G)A = FA + GA, \tag{22}$$

$$(FG)A = F(GA). \tag{23}$$

By using the commutative law of addition in (1), it is easily shown that the addition of two operators is commutative, so that $F + G = G + F$. On the other hand, operator multiplication is in general non-commutative, so that

$$FG \neq GF, \tag{24}$$

and, in the exceptional cases when $FG = GF$, we will say that the two operators F and G commute.

Powers of a linear operator F are defined by a series of repeated multiplications according to (23):

$$F^2 = FF, \quad F^3 = F.F^2, \ldots \quad F^{m+1} = FF^m, \tag{25}$$

and they may then be used to define polynomials of an operator:

$$P(F) = a_0 + a_1 F + a_2 F^2 + \cdots + a_m F^m, \tag{26}$$

where a_o, a_1, a_2, ... a_n are complex constants. It is easy to prove that any polynomial operator P(F) is a linear operator, if F is a linear operator.

<u>Inverse operators.</u> - Let us now introduce a new concept connected with the inverse of the mapping $A \longrightarrow B$, i.e. the mapping $B \longrightarrow A$. If T is a linear operator such that there exists a <u>unique</u> element A in the domain of T corresponding to any given element B in the range of T according to the relation B = TA, then there exists a unique mapping of B on A, and the associated operator is called the inverse of T and is denoted by T^{-1}:

$$A = T^{-1} B \tag{27}$$

It is easily shown that, if T is a linear operator, then the inverse T^{-1} is also a linear operator.

According to the definition, one has to show that every element B has a unique "image element" A to see that T^{-1} exists. A considerable simplification is hence rendered by the fact that it is actually sufficient to check that this happens for the single element $B = \overline{0}$, according to the following theorem:

> The operator T^{-1} exists, if and only if the relation $TA = \overline{0}$ implies $A = \overline{0}$. (28)

Before making the proof, we observe that every linear operator maps the zero-element of its domain on the zero-element of its range, since $T\bar{0} = T(0 \cdot A) = 0 \cdot TA = \bar{0}$. Let us first assume that T^{-1} exists. Since the mapping is now unique, the image element $B = \bar{0}$ corresponds to $A = \bar{0}$, i.e. $TA = \bar{0}$ implies $\bar{A} = 0$, which proves the first part of the theorem.

In order to prove the second part, one starts from the assumption that $TA = \bar{0}$ implies $A = \bar{0}$. It is easy to see that the inverse mapping must be unique for, if there would be two elements A' and A" corresponding to one and the same image element B, one would have

$$\begin{aligned} TA &= B, \qquad TA'' = B, \\ T(A' - A'') &= \bar{0}, \\ A' - A'' &= \bar{0}, \\ A' &= A'' \end{aligned} \tag{29}$$

i.e. one would obtain a contradiction. Hence, the inverse mapping is unique, and T^{-1} exists.

<u>Matrix representations of operators.</u> - Let us now consider a finite space of order n which has a basis $\mathbb{X} = (X_1, X_2, \ldots X_n)$. According to (5), every element A may be expressed in the form

$$A = \mathbb{X}\,a = \sum_{k=1}^{n} X_k a_k, \tag{30}$$

where the coefficients a_k are unique and denoted by the symbol $a_k = X_k, \mathbb{X}|A$. A linear operator T is assumed to map the elements of the linear space onto itself or onto a subspace of itself, and TA is hence an element of the space which may be expressed in terms of the basis $\mathbb{X}$. In order to treat this problem, we will introduce the image elements of the elements X_l of the basis through the relations

$$T X_l = \sum_{k=1}^{n} X_k T_{kl} \tag{31}$$

where the complex numbers $T_{k\ell}$ are the uniquely determined expansion coefficients given by the symbol:

$$T_{k\ell} = \{ X_k , X \mid T X_\ell \} \tag{32}$$

Using the properties of linear operators, one obtains from (30) the formula

$$T A = \sum_{k,\ell=1}^{n} X_k T_{k\ell} a_\ell \tag{33}$$

which shows that a linear operator T is fully characterized by the n^2 complex numbers $T_{k\ell}$. It is convenient to arrange these numbers for $k, \ell = 1, 2, 3, \ldots n$ into a square matrix $\mathbf{T}_X$:

$$\mathbf{T}_X = \begin{pmatrix} T_{11} & T_{12} & T_{13} & \ldots & T_{1n} \\ T_{21} & T_{22} & T_{23} & \ldots & T_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ T_{n1} & T_{n2} & T_{n3} & \ldots & T_{nn} \end{pmatrix} , \tag{34}$$

which will be called the <u>matrix representation</u> of the linear operator T with respect to the basis $\mathbf{X}$.

We note that the symbol (32) does not give us any recipe for the evaluation of the matrix elements, and that this finally depends on the realization of the elements of the space. The matrix $\mathbf{T}_X$ is here solely defined through the relations (31), which may be condensed into the matrix formula

$$T \mathbf{X} = \mathbf{X} \mathbf{T}_X \tag{35}$$

The expansion theorem $A = \mathbf{X} a$ gives then directly

$$T A = \mathbf{X} \mathbf{T}_X \mathbf{a} , \tag{36}$$

which is the matrix form for the general formula (33).

The sum and product for two operators F and G were defined by the relations (22) and (23), respectively. For the matrix representation of the sum, one has the rule

$$(F+G)_{kl} = F_{kl} + G_{kl} . \tag{37}$$

The proof follows from the fact that

$$\begin{aligned} (F+G)\mathbf{X} &= F\mathbf{X} + G\mathbf{X} = \\ &= \mathbf{X}\mathbf{F}_X + \mathbf{X}\mathbf{G}_X = \mathbf{X}(\mathbf{F}_X + \mathbf{G}_X) \end{aligned} \tag{38}$$

For the product, one obtains similarly

$$\begin{aligned} (FG)\mathbf{X} &= F(G\mathbf{X}) = F(\mathbf{X}\mathbf{G}_X) = \\ &= (F\mathbf{X})\mathbf{G}_X = (\mathbf{X}\mathbf{F}_X)\mathbf{G}_X = \\ &= \mathbf{X}(\mathbf{F}_X\mathbf{G}_X) , \end{aligned} \tag{39}$$

which shows that the matrix of an operator product is the matrix product of the matrices of the individual factors. According to (9), one has for each element

$$(FG)_{kl} = \sum_\alpha F_{k\alpha} G_{\alpha l} . \tag{40}$$

Using these rules, one can now prove that every algebraic relation between operators corresponds to a similar algebraic relation between the matrix representations. We note particularly that, if the operator T has an inverse T^{-1}, the latter operator has a matrix representation given by the inverse matrix $\mathbf{T}_x^{-1}$.

Similarity transformations. - Let us finally consider the transformation of a matrix representation under a change of basis. If $\mathbf{X} = (X_1, X_2, \ldots X_n)$ is one basis and $\mathbf{Y} = (Y_1, Y_2, \ldots Y_n)$ is another, one has according to (14) and (17) the connections

$$\mathbf{Y} = \mathbf{X}\alpha \ , \qquad \mathbf{X} = \mathbf{Y}\beta \tag{41}$$

where $\beta = \alpha^{-1}$. The matrix representations $\mathbf{T}_X$ and $\mathbf{T}_Y$ of a linear operator T are further defined by the relations

$$T\mathbf{X} - \mathbf{X}\,\mathbf{T}_X \ , \qquad T\mathbf{Y} = \mathbf{Y}\,\mathbf{T}_Y \ , \tag{42}$$

respectively. This gives immediately

$$\mathbf{Y}\,\mathbf{T}_Y = T\mathbf{Y} = T\mathbf{X}\alpha = \mathbf{X}\,\mathbf{T}_X\alpha =$$
$$= \mathbf{Y}\beta\,\mathbf{T}_X\alpha = \mathbf{Y}\,(\beta\,\mathbf{T}_X\alpha) \ ,$$

$$\mathbf{Y}\,(\mathbf{T}_Y - \beta\,\mathbf{T}_X\alpha) = \bar{0} \ ,$$
$$\mathbf{T}_Y - \beta\,\mathbf{T}_X\alpha = 0 \ , \tag{43}$$

since $\mathbf{Y}$ is a linearly independent subset. Under a change of basis (41), one obtains hence the following transformation formulas

$$\mathbf{T}_Y = \beta\,\mathbf{T}_X\alpha \ , \qquad \mathbf{T}_X = \alpha\,\mathbf{T}_Y\beta \ , \tag{44}$$

with $\beta = \alpha^{-1}$, which are called <u>similarity transformations</u>.

<u>Projection Operators.</u> - Starting from the expansion theorem (30) in the form $A = \sum_k X_k a_k$, we will now consider the operators O_k defined by the relation

$$O_k A = X_k a_k \ ; \tag{45}$$

for $k = 1, 2, 3, \ldots n$. This implies that the operator O_k maps an element A onto its component $X_k a_k$, or that it selects the k^{th} component out of the expansion. Using the definitions, it is easily shown that

$$\mathcal{O}_k(A_1 + A_2) = \mathcal{O}_k A_1 + \mathcal{O}_k A_2 \ , \quad \mathcal{O}_k(\alpha A) = \alpha \mathcal{O}_k A \ , \tag{46}$$

which means that O_k is a linear operator. Since the repeated use of O_k, i.e. the selection of the k^{th} component out of the k^{th} component, still leads to the same result, one has $O_k^2 = O_k$. One says that the operator O_k is "idempotent" and, for geometrical reasons, one speaks also of a <u>projection operator</u>. This concept is defined in various ways in different parts of the literature, but here we will use the terms idempotent operators and projection operators as synonymous. Since the selection of the k^{th} component out of the ℓ^{th} component for $k \neq \ell$ necessarily gives a zero-element, one has further the operator relation $O_k O_\ell = 0$ and says that the operators O_k and O_ℓ are "mutually exclusive". In summary, we have hence

$$\mathcal{O}_k^2 = \mathcal{O}_k \ , \qquad \mathcal{O}_k \mathcal{O}_\ell = 0 \ , \quad k \neq \ell \ . \tag{47}$$

Using the expansion theorem (5) and (22), one can further see that

$$\begin{aligned} A &= \sum_{k=1}^{m} \mathcal{O}_k A = \Big(\sum_{k=1}^{m} \mathcal{O}_k \Big) A \ , \\ \Big(I - \sum_{k=1}^{m} \mathcal{O}_k \Big) A &= \bar{0} \ , \end{aligned} \tag{48}$$

for every element A, which shows that the operator $(I - \sum_k \mathcal{O}_k)$ must be a zero-operator. Hence one has the relation

$$I = \sum_{k=1}^{m} \mathcal{O}_k \ . \tag{49}$$

The operators $O_1, O_2, \ldots O_n$ form a family of mutually exclusive projection operators, which together adds up to the identity operators. One says also that relation (49) is a "resolution of the identity" in terms of projection operators.

Let us now consider the operators Q_p which are defined by the relation

$$Q_p = \sum_{k=1}^{p} O_k . \tag{50}$$

Using the two relations (47), one finds immediately that $Qp^2 = Qp$, i.e. Q_p is also idempotent. One gets particularly

$$Q_p A = \sum_{k=1}^{p} X_k a_k , \tag{51}$$

which is an element belonging to the linear manifold spanned by the subset $X_1, X_2, \ldots X_p$. One says that the element Q_pA is the "projection" of A with respect to this manifold out of the basis $\mathbf{X} = (X_1, X_2, \ldots X_n)$, and Q_p is the associated projection operator.

By using the notation (12), one finds that $O_kA = X_k a_k = X_k\{X_k, \mathbf{X}|A\}$ for every element A, and it is hence suggestive to try to write the projection operator symbolically in the form

$$O_k = X_k \{X_k, \mathbf{X}| . \tag{52}$$

For the projection on a subspace of order p, this gives

$$Q_p = \sum_{k=1}^{p} X_k \{X_k, \mathbf{X}| , \tag{53}$$

and, for the resolution of the identity, one obtains particularly $I = \sum_k X_k\{X_k, \mathbf{X}|$. These notations are here of a purely formal nature, but they will later turn out to be quite forceful.

In conclusion, we will study the matrix representations of the operators O_k according to (31). Using (45), we obtain

$$O_k X_k = X_k ; \qquad O_k X_l = 0 , \ k \neq l , \tag{54}$$

showing that the matrix representation of O_k has a single non-zero element, which equals 1 and is placed in the k^{th} position of the diagonal so that

$$\mathbf{0}_1 = \begin{pmatrix} 1\,0\,0 \cdots \\ 0\,0\,0 \cdots \\ 0\,0\,0 \cdots \\ \cdots\cdots \end{pmatrix} ; \qquad \mathbf{0}_2 = \begin{pmatrix} 0\,0\,0 \cdots \\ 0\,1\,0 \cdots \\ 0\,0\,0 \cdots \\ \cdots\cdots \end{pmatrix} ; \tag{55}$$

It is easily checked that these matrices satisfy the fundamental algebraic relations (47) and (49).

Trace of an operator. - The "trace" of a quadratic matrix is defined as the sum of the diagonal elements:

$$\mathrm{Tr}\,(\mathbf{M}) = \sum_k M_{kk} . \tag{56}$$

If $\mathbf{M} = \mathbf{K} \cdot \mathbf{L}$, where $\mathbf{K}$ and $\mathbf{L}$ are two quadratic or compatible rectangular matrices, one has the theorem

$$\mathrm{Tr}\,(\mathbf{K} \cdot \mathbf{L}) = \mathrm{Tr}\,(\mathbf{L}\,\mathbf{K}) , \tag{57}$$

even if the two matrices in general do not commute. This depends on the fact that

$$\begin{aligned} \mathrm{Tr}\,(\mathbf{K}\cdot\mathbf{L}) &= \sum_k (\mathbf{K}\cdot\mathbf{L})_{kk} = \sum_k \Big(\sum_\alpha K_{k\alpha} L_{\alpha k}\Big) = \\ &= \sum_\alpha \Big(\sum_k L_{\alpha k} K_{k\alpha}\Big) = \sum_\alpha (\mathbf{L}\cdot\mathbf{K})_{\alpha\alpha} = \mathrm{Tr}\,(\mathbf{L}\cdot\mathbf{K}) . \end{aligned} \tag{58}$$

Using (57), one can immediately prove that the trace of quadratic matrix which is a product of a finite number of quadratic or compatible rectangular matrices is invariant under a cyclic permutation of the factors.

The trace of an operator T is defined as the trace of one of its matrix representations:

$$\mathrm{Tr}\,(\mathrm{T}) = \mathrm{Tr}\,(\mathbf{T}_X)\;, \tag{59}$$

and we note that this quantity is independent of the choice of the representation. According to (44) and (58), one has

$$\begin{aligned}\mathrm{Tr}\,(\mathbf{T}_Y) &= \mathrm{Tr}\,(\boldsymbol{\beta}\,\mathbf{T}_X\,\boldsymbol{\alpha}) = \mathrm{Tr}\,(\mathbf{T}_X\,\boldsymbol{\alpha}\,\boldsymbol{\beta}) = \\ &= \mathrm{Tr}\,(\mathbf{T}_X)\;,\end{aligned} \tag{60}$$

which proves our statement. Simple examples are provided by the projection operators. From (55) follows directly

$$\mathrm{Tr}\,(\mathbb{O}_{k}) = 1 \tag{61}$$

whereas one has $\mathrm{Tr}(Q_p) = p$ according to (50).

3. EIGENVALUE PROBLEM

Let us consider a linear mapping of a given linear space represented by the operator T. The problem is whether there are any non-zero elements C forming "points" or "lines" which are invariant under the transformation

$$T C = \lambda C \tag{62}$$

This is an eigenvalue problem, and the non-trivial solutions C are called eigenelements and the constant λ the associated eigenvalue. Geometrically the eigenvalue problem is connected with the question of finding the "rotation axis" of the transformation, and it is sometimes also called the "pole problem". Equation (62) is of fundamental importance not only for quantum theory but for large parts of mathematics and physics in general.

The eigenvalue problem may be given an alternative formulation. From (62) follows that $(T - \lambda \cdot I) C = \bar{0}$ for $C \neq \bar{0}$, and, according to (28), this implies that the operator $(T - \lambda \cdot I)$ has no inverse, i.e. that the operator $(T - \lambda \cdot I)^{-1}$ becomes singular for the eigenvalues. In many connections, it is convenient to introduce the "resolvent" of T, which is the inverse operator $(T - z \cdot I)^{-1}$, where z is a complex variable. We note that the resolvent exists for all values of z, except the eigenvalues λ . It is possible to develop the entire eigenvalue theory on this basis.

Here we will instead proceed in another way based on the use of the matrix representations. Considering a linear space of order n, we will introduce a basis $\mathbf{X} = (X_1, X_2, \ldots X_n)$ and the matrix representation $\mathbf{T}_x$ of T defined by (35). Expansion of the eigenelement C gives

$$C = \sum_k X_k c_k = \mathbf{X}\mathbf{c} \quad , \tag{63}$$

where the coefficients c_k form a column vector $\mathbf{c}$. The eigenvalue problem (62) may be written in the form $(T - \lambda \cdot I)\, C = \bar{0}$, and this leads to the matrix relation

$$
\begin{aligned}
(T - \lambda \cdot I)\, C &= (T - \lambda \cdot I)\, \mathbf{X} \mathbf{c} = \\
&= \mathbf{X}\,(\mathbf{T}_X - \lambda \cdot \mathbf{1})\,\mathbf{c} = \bar{0}
\end{aligned}
\tag{64}
$$

However, since the set $\mathbf{X}$ is assumed to be linearly independent, every relation $\mathbf{X}\mathbf{a} = \bar{0}$ implies $\mathbf{a} = \mathbf{0}$, where $\mathbf{0}$ is a column vector with the elements 0, and hence we obtain

$$
\boxed{(\mathbf{T}_X - \lambda \cdot \mathbf{1})\,\mathbf{c} = \mathbf{0}\ .}
\tag{65}
$$

This is the matrix form of the eigenvalue problem (62), and it is equivalent with a homogeneous system of linear equations:

$$
\sum_{l=1}^{m} (T_{kl} - \lambda \cdot \delta_{kl})\, c_l = 0 \ ,
\tag{66}
$$

for k = 1, 2, ... n. Such a system has a non-trivial solution, if and only if

$$
\det \{ T_{kl} - \lambda \cdot \delta_{kl} \} = 0 \ ,
\tag{67}
$$

which is the well-known "secular equation".

The equations (63), (66), and (67) form the basis for a large part of quantum chemistry, and good examples are provided by the MO-LCAO-method and the method using "superposition of configurations". However, since there is no scalar product introduced here, there cannot be any non-orthogonality problem connected with (66), and this indicates that the matrix elements T_{kl} defined by (31) may have a somewhat different meaning than usual; this problem will be studied in greater detail in a later section.

We note that the matrix equation (65) is "covariant" under a change of basis, say $\mathbf{Y} = \mathbf{X}\boldsymbol{\alpha}$. According to (44) and (18), one has

$$T_Y = \beta T_X \alpha , \qquad c_Y = \beta c_X , \tag{68}$$

where $\beta = \alpha^{-1}$, and hence, we obtain

$$(T_Y - \lambda \cdot \mathbf{1}) c_Y = \beta (T_X - \lambda \cdot \mathbf{1}) c_X = 0 ,$$

which proves our statement.

Characteristic polynomial. - Let us now define a function of the complex variable z by the relation:

$$P(z) = \det \{ T_{kl} - z \delta_{kl} \} . \tag{69}$$

It is easily seen that $P(z)$ is a polynomial of degree n:

$$P(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_{n-1} z^{n-1} + a_n z^n , \tag{70}$$

where $a_n = (-1)^n$, and $P(z)$ is called the "characteristic polynomial" associated with the linear operator T. The coefficients are independent of the choice of representation, since one has

$$\begin{aligned} \det\{\mathbf{T}_Y - z\cdot\mathbf{1}\} &= \det\{\beta(\mathbf{T}_X - z\cdot\mathbf{1})\alpha\} = \\ &= \det\beta\cdot\det\{\mathbf{T}_X - z\cdot\mathbf{1}\}\cdot\det\alpha = \\ &= \det\{\mathbf{T}_X - z\cdot\mathbf{1}\} , \end{aligned} \tag{71}$$

where the last simplification is obtained by using the fact that $\det\beta\cdot\det\alpha = = \det(\beta\alpha) = \det(\mathbf{1}) = 1$.

The characteristic polynomial has exactly n roots in the complex plane $\lambda_1, \lambda_2, \ldots \lambda_n$ which are the eigenvalues of the problem. The factorial theorem gives immediately

$$P(z) = \prod_{k=1}^{n} (\lambda_k - z) \tag{72}$$

There may be multiple roots λ_k, and the degree of multiplicity g_k is also called the "order of degeneracy" of the eigenvalue. An eigenvalue is "non-degenerate" if it has $g = 1$, i.e. if the root is distinct. In the following, we will first consider the case of all roots distinct, and later we will study the general case.

<u>Case of all roots distinct.</u> - In this case, all the roots λ_k are single roots, and one has n distinct eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$ in the complex plane. The eigenvalue problem (62) takes the form $\mathbf{T}C_k = \lambda_k C_k$ or

$$(T - \lambda_k \cdot I)\, C_k = \bar{0} \tag{73}$$

for $k = 1, 2, \ldots n$. We see that the eigenelement C_k is eliminated by the operator $(T - \lambda_k \cdot I)$, which in this connection will be called an "eliminator". The following theorem is of fundamental importance:

> If all roots $\lambda_1, \lambda_2, \ldots \lambda_m$ of the secular equation are distinct, the associated set of eigen-elements $C_1, C_2, \ldots C_n$ are linearly independent. (74)

The proof is simple. Let us consider a linear relation of the type

$$\sum_{k=1}^{m} C_k \alpha_k = \bar{0} ,$$

and let us operate on this equation with the product of the "eliminators" for $k = 2, 3, \ldots n$, i.e. with

$$\prod_{l=2}^{m} (T - \lambda_l \cdot I) \tag{75}$$

which gives $\prod_{l=2}^{m} (\lambda_1 - \lambda_l)\, C_1 \alpha_1 = \bar{0}$, and hence $\alpha_1 = 0$. In a similar way, one shows that $\alpha_2 = \alpha_3 = \ldots = \alpha_m = 0$ by using eliminator-products which will let only the term for $k = 2, 3, \ldots n$, respectively, survive. The theorem is thus proven.

Since one has now a set of n linearly independent elements $\mathbf{C} = (C_1, C_2, \ldots C_n)$ in a space of the same order, one knows according to (13) that this set forms a basis of the space. This gives the theorem about expansions in eigenelements:

If all the roots of the secular equation are distinct, one may expand an arbitrary element A of the linear space in terms of the eigenelements $C_1, C_2, \ldots C_n$ of the operator T:

$$A = \sum_k C_k a_k \tag{76}$$

Let us now choose the set $\mathbf{C} = (C_1, C_2, \ldots C_n)$ as the basis for the matrix representation of T. The eigenvalue relation $T C_k = C_k \lambda_k$ may, according to (31), be interpreted so that $\mathbf{T}_c$ has the special form:

$$\mathbf{T}_c = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & \cdot \\ 0 & \lambda_2 & 0 & 0 & \vdots \\ 0 & 0 & \lambda_3 & 0 & \vdots \\ 0 & 0 & 0 & \lambda_4 & \vdots \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} = \boldsymbol{\lambda} \tag{77}$$

which is called a "diagonal matrix". Using (44), we can then say that there exists a similarity transformation which brings the matrix $\mathbf{T}_x$ to diagonal form. According to (63), one has $C_k = \mathbf{X} \mathbf{c}_k$ and, arranging the column vectors $\mathbf{c}_k$ in a row, one obtains a quadratic matrix:

$$\boldsymbol{\gamma} = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \ldots \mathbf{c}_m) = \begin{pmatrix} c_{11} & c_{12} & c_{13} & \ldots & c_{1m} \\ c_{21} & c_{22} & c_{23} & \ldots & c_{2m} \\ c_{31} & c_{32} & c_{33} & \ldots & c_{3m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{m1} & c_{m2} & c_{m3} & \ldots & c_{mm} \end{pmatrix} \tag{78}$$

and the transformation $\mathbf{C} = \mathbf{X} \boldsymbol{\gamma}$ This gives

$$\boldsymbol{\gamma}^{-1} \mathbf{T}_X \boldsymbol{\gamma} = \boldsymbol{\lambda} \tag{79}$$

showing that, in the distinct case, the matrix $\mathbf{T}_x$ can always be brought to diagonal form.

Cayley - Hamilton theorem. - A polynomial of an operator T is defined by the expression (26). For the characteristic polynomial $P(z)$ defined by (70), one has the Cayley-Hamilton theorem:

$$\boxed{P(T) = 0\ ,} \tag{80}$$

i.e. $P(T)$ is identical to the zero-operator. In the distinct case, the proof is simple. According to (76), an arbitrary element A of the space may be expanded in the form

$$A = \sum_{k=1}^{m} C_k a_k \tag{81}$$

However, relation (72) shows that one has also $P(T) = \prod_{k=1}^{m} (\lambda_k I - T)$, i.e. $P(T)$ is a product of all eliminators, so that

$$P(T) A = \sum_{k=1}^{m} P(\lambda_k) C_k a_k = \breve{0} \tag{82}$$

Since this happens for every A, one has $P(T) = 0$.

In the general case of roots of various multiplicities, we will start with the matrix relation (16) applied to the operator $\mathbf{M} = \mathbf{T}_x - z\,\mathbf{1}$. Letting $\mathbf{N}$ be the matrix of the cofactors N_{kl} to the element M_{kl} in $\mathbf{M}$, we note that each element N_{kl} is a polynomial of degree $(n-1)$ in the variable z, and that $\mathbf{N}$ hence may be written in the form:

$$\mathbf{N} = \mathbf{N}_0 + \mathbf{N}_1 z + \mathbf{N}_2 z^2 + \dots \qquad + \mathbf{N}_{n-1} z^{n-1}\ , \tag{83}$$

where $\mathbf{N}_p$ is the matrix of the coefficients for z^p. Application of (16) gives immediately:

$$\mathbf{M} \cdot \widetilde{\mathbf{N}} = \det\{\mathbf{M}\} \cdot \mathbf{1}\ , \tag{84}$$

or

$$(\mathbf{T} - z \cdot \mathbf{1})(\widetilde{\mathbf{N}}_0 + \widetilde{\mathbf{N}}_1 z + \dots + \widetilde{\mathbf{N}}_{n-1} z^{n-1}) = P(z) \cdot \mathbf{1} \tag{85}$$

Multiplying together and separating after powers of z, one obtains

$$\left\{\begin{aligned} \mathbf{T}\tilde{\mathbf{N}}_0 &= a_0 \cdot \mathbf{1} \,, \\ \mathbf{T}\tilde{\mathbf{N}}_1 - \tilde{\mathbf{N}}_0 &= a_1 \cdot \mathbf{1} \,, \\ \mathbf{T}\tilde{\mathbf{N}}_2 - \tilde{\mathbf{N}}_1 &= a_2 \cdot \mathbf{1} \,, \\ \cdots\cdots & \cdots\cdots \\ \mathbf{T}\tilde{\mathbf{N}}_{m-1} - \tilde{\mathbf{N}}_{m-2} &= a_{m-1} \cdot \mathbf{1} \,, \\ -\tilde{\mathbf{N}}_{m-1} &= a_m \cdot \mathbf{1} \,. \end{aligned}\right. \tag{86}$$

By multiplying these equations successively by $\mathbf{1}$, $\mathbf{T}$, $\mathbf{T}^2$, ... $\mathbf{T}^n$ and by summing all of them together, one gets finally

$$\mathbf{0} = a_0 \cdot \mathbf{1} + a_1 \cdot \mathbf{T} + a_2 \cdot \mathbf{T}^2 + \cdots + a_m \mathbf{T}^m \,, \tag{87}$$

which proves the Cayley-Hamilton theorem $P(\mathbf{T}) = \mathbf{0}$ in matrix form in the general case. This gives immediately $P(T) = 0$ by means of the equivalence theorem previously discussed.

4. PROJECTION OPERATORS AS EIGENOPERATORS; RESOLUTION OF IDENTITY IN THE CASE OF DISTINCT EIGENVALUES

Eigenprojectors. - Let us again consider a linear space of finite order n. In a previous section, it has been shown that, with every basis $\mathbf{X} = (X_1, X_2, \dots X_n)$, there is associated a family of projection operators $O_1, O_2, \dots O_n$ which are idempotent, mutually exclusive and form a resolution of the identity according to (47) and (49). Let us now particularly study those projection operators $O_1, O_2, \dots O_n$ as are associated with the eigenbasis $\mathbf{C} = (C_1, C_2, \dots C_n)$ to a linear operator T having only distinct eigenvalues λ_k . One may write the expansion theorem (81) in the form

$$A = \sum_{k=1}^{n} A_k \,, \tag{88}$$

where $A_k = C_k a_k$, and one may consider (88) as an "analysis" of an element A in terms of eigenelements to T, satisfying the relation

$$T A_k = \lambda_k A_k \tag{89}$$

According to (45), the projection operator O_k is defined through the selection property:

$$O_k A = A_k \,, \tag{90}$$

and one says that the term or "component" A_k is the projection of the element A on the eigenspace of T associated with the eigenvalue λ_k Since one has $(T O_k - \lambda_k O_k) A = \overline{0}$ for an arbitrary element A, the operator $(T O_k - \lambda_k O_k)$ is necessarily a zero-operator, which gives

$$T O_k = \lambda_k O_k \tag{91}$$

The projection operator O_k satisfies hence the fundamental eigenvalue relation (62), and O_k may be characterized as an "eigenoperator" or "eigenprojector" to T. According to (47) and (49), the projection operators satisfy further the basic formulas

$$\mathbb{O}_k^2 = \mathbb{O}_k ; \qquad \mathbb{O}_k \mathbb{O}_l = 0, \quad k \neq l \tag{92}$$

$$I = \sum_{k=1}^{m} \mathbb{O}_k \tag{93}$$

Letting the operator T work on both sides of (93) and using (91), one obtains

$$T = \sum_{k=1}^{m} \lambda_k \mathbb{O}_k , \tag{94}$$

which is called the "spectral resolution" of the operator T. If f(z) is an arbitrary polynomial in the complex variable z, one gets further

$$f(T) = \sum_{k=1}^{m} f(\lambda_k) \mathbb{O}_k , \tag{95}$$

and, from the polynomial, one can then proceed to define an arbitrary algebraic function of T.

It is clear that we are here treating an almost trivial case, but the important thing is that all the concepts introduced are of fundamental character and may be generalized.

<u>Matrix representation of the eigenprojectors.</u> - In the $\mathbf{C}$-basis the eigenprojectors O_k have matrix representations of the form (55), i.e.

$$\mathbf{O}_{k,C} = \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \ddots\, 1 \ddots & \\ & & & 0 \end{bmatrix} \tag{96}$$

i.e. there is a single non-zero element, which equals 1 and is placed in the k^{th} position of the diagonal. It is now possible to derive the matrix representations in the $\mathbf{X}$-basis by means of the general transformation formulas (44). One has $\mathbf{C} = \mathbf{X}\gamma$ and $\mathbf{X} = \mathbf{C}\gamma^{-1}$, where $\gamma = (\mathbf{c}_1, \mathbf{c}_2, \ldots \mathbf{c}_m)$ is specified in (78). It is convenient to write out the matrix γ^{-1} explicitly in the form:

$$\gamma^{-1} = \begin{pmatrix} d_{11} & d_{12} & \ldots & d_{1m} \\ d_{21} & d_{22} & \ldots & d_{2m} \\ \cdot & \cdot & & \cdot \\ d_{m1} & d_{m2} & \ldots & d_{mm} \end{pmatrix} = \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_m \end{pmatrix} \tag{97}$$

where $\mathbf{d}_1, \mathbf{d}_2, \ldots \mathbf{d}_n$ stand for a set of row vectors. Transforming the matrix (96) to the $\mathbf{X}$-basis, one obtains

$$\mathbf{O}_k = \gamma \begin{pmatrix} 0 & & & \\ & 0 \ddots & & \\ & & 1 & \\ & & & \ddots 0 \end{pmatrix} \gamma^{-1} = \mathbf{c}_k \mathbf{d}_k \tag{98}$$

One has further $\gamma^{-1}\gamma = \gamma\gamma^{-1} = \mathbf{1}$, which gives

$$\mathbf{d}_k \mathbf{c}_k = \delta_{kl}, \qquad \mathbf{1} = \sum_k \mathbf{c}_k \mathbf{d}_k \tag{99}$$

We note that the second relation is nothing but the "resolution of the identity" (93), and that the "spectral resolution" of the matrix $\mathbf{T}_X$ in analogy to (94) takes the form

$$\mathbf{T}_X = \sum_{k=1}^{m} \lambda_k \mathbf{O}_k = \sum_{k=1}^{m} \lambda_k \mathbf{c}_k \mathbf{d}_k \tag{100}$$

This formula gives a simple recipe for constructing a matrix with given distinct eigenvalues λ_k and eigenvectors $\mathbf{c}_k$. One should first combine the vectors $\mathbf{c}_k$ to a quadratic matrix γ according to (78), and, after evaluation of the inverse γ^{-1}, one obtains the row vectors $\mathbf{d}_k$. Multiplication of the column vector $\mathbf{c}_k$ and the row vector $\mathbf{d}_k$ gives the matrix $\mathbf{O}_k$, which may then be combined with an arbitrarily chosen eigenvalue λ_k. Summation of the various contributions according to (100) gives finally the matrix desired.

<u>Product form for the eigenprojectors.</u> - In this section, we will arrive to the concept of the eigenprojectors in a completely different way, which has certain advantages in the generalizations to be carried out later. Again we will consider a linear space of finite order n and a linear operator T having only distinct eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_m$ with the eigenelements $C_1, C_2, C_3, . \quad C_n$. In the proof for the linear independence of this set, we used operators of the type (75) which are products of "eliminators" according to (73). One has particularly

$$\left\{ \prod_{j \neq k} (T - \lambda_j \cdot I) \right\} C_\ell = \left\{ \prod_{j \neq k} (\lambda_\ell - \lambda_j) \right\} C_\ell , \tag{101}$$

which shows that the right-hand side will vanish except for $k = \ell$ This gives further

$$\left\{ \prod_{j \neq k} \frac{T - \lambda_j \cdot I}{\lambda_k - \lambda_j} \right\} C_\ell = C_k \, \delta_{k\ell} \tag{102}$$

The product operator in the left-hand member has hence exactly the same characteristic property (54) as the projection operator O_k previously introduced, and one obtains the alternative form

$$O_k = \prod_{j \neq k} \frac{T - \lambda_j \cdot I}{\lambda_k - \lambda_j} \tag{103}$$

Using the Cayley-Hamilton theorem in the product form $P(T) = \prod_k (\lambda_k \cdot I - T)$, one can now easily give alternative proofs for the fundamental relations (91)-(94), and we will return to this approach in a later section in connection with the generalization to linear spaces of an <u>infinite</u> order

5. CLASSICAL CANONICAL FORM OF A MATRIX OF FINITE ORDER

Nilpotent operators. - Before starting the general treatment, we will consider a special class of operators called "nil-potent operators" with certain fundamental properties:

An operator N is said to be nilpotent of order p, if

$$N^p = 0, \tag{104}$$

and $N^{p-1} \neq 0$.

In order to study such an operator, we will introduce a certain basis. Since $N^{p-1} \neq 0$, there exists at least one element $D_p \neq \overline{0}$, such that $N^{p-1} D_p \neq \overline{0}$. Starting from this element, one can now define a series of elements $D_{p-1}, D_{p-2}, \ldots D_2, D_1$ successively through the relations

$$\begin{aligned} D_{p-1} &= N D_p, \\ D_{p-2} &= N D_{p-1} = N^2 D_p, \\ D_{p-3} &= N D_{p-2} = N^2 D_{p-1} = N^3 D_p, \\ &\cdots\cdots\cdots \\ D_1 &= N D_2 = \cdots = N^{p-2} D_{p-1} = N^{p-1} D_p \neq \overline{0} \end{aligned} \tag{105}$$

They are all different from the zero-element, and they fulfill the relations $ND_1 = N^2 D_2 = N^3 D_3 = \ldots = N^{p-1} D_{p-1} = \overline{0}$. It is now easily seen that the elements $D_1, D_2, D_3, \ldots D_p$ are linearly independent. Starting from the linear relation

$$D_1 \alpha_1 + D_2 \alpha_2 + \cdots + D_{p-1} \alpha_{p-1} + D_p \alpha_p = \overline{0} \tag{106}$$

and multiplying to the left by N^{p-1}, one obtains $D_1 \alpha_p = \overline{0}$, i.e. $\alpha_p = 0$. Substitution into (106) gives the simplified relation

$$D_1 \alpha_1 + D_2 \alpha_2 + \cdots + D_{p-1} \alpha_{p-1} = \overline{0}, \tag{107}$$

and multiplication to the left by N^{p-2} gives further $D_1 \alpha_{p-1} = \bar{0}$, i.e. $\alpha_{p-1} = 0$. Proceeding in the same way, one obtains finally $\alpha_1 = \alpha_2 = \dots = \alpha_p = 0$, which shows that the set $D_1, D_2, \dots D_p$ is linearly independent.

Let us now first consider the case $p = n$, where n is the order of the space. One can now choose the set $\mathbf{D} = (D_1, D_2, \dots D_n)$ as a basis. From (105) follows that

$$\begin{cases} N D_1 = \bar{0}, \\ N D_2 = D_1, \\ N D_3 = D_2, \\ \dots\dots \\ N D_{n-1} = D_{n-2}, \\ N D_n = D_{n-1} \end{cases} \tag{108}$$

Interpreting these equations according to (31), one obtains the following matrix representation of N in the $\mathbf{D}$-basis:

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & & & & & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \tag{109}$$

This matrix has zeros everywhere, except in the first diagonal above the main diagonal. A matrix of this type is called a "Jordan block", and it represents the classical canonical form of a nilpotent matrix. The Jordan blocks of order 1, 2, 3, 4, ... take the special form

$$\begin{bmatrix} 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \dots \tag{110}$$

and we note that such a matrix can never be brought to diagonal form by a similarity transformation. Nilpotent matrices are of fundamental importance in physics in connection with so-called shift-operators, for instance, the step-up and step-down operators M_+ and M_- in the theory of angular momenta.

In the case $p < n$, we observe first that every matrix of order n which consists of a diagonal series of Jordan blocks is necessarily nilpotent of an order which equals the order of the largest Jordan block. For the case $n = 4$ one has, for instance, the following possibilities

$$\underset{(4)}{\overset{p=4}{\begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}}} \quad \underset{(3,1)}{\overset{p=3}{\begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ & & & 0 \end{bmatrix}}} \quad \underset{(2,2)}{\overset{p=2}{\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}}} \quad \underset{(2,1,1)}{\overset{p=2}{\begin{bmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}}} \quad \underset{(1,1,1,1)}{\overset{p=1}{\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}}} \tag{111}$$

Below the matrices, the orders of the Jordan blocks entering the entire matrix are indicated, and these numbers are called the "Segré characteristics" of the matrix. The number of types occurring corresponds to the number of partitionings of the integer n.

In order to prove that every nilpotent matrix having $p < n$ may be written in this special form, we will consider the subspace V_p of the original space V which is such that, for every element A in V_p, one has $N^{p-1}A \neq 0$. Let the order of the subspace V_p be q, and let us span this space by means of a linearly independent set of elements D_p', D_p'', ... $D_p^{(q)}$ For every set of complex numbers $\alpha_p', \alpha_p'', \dots \alpha_p^{(q)}$ which are not all identically to zero, one has consequently

$$N^{p-1}\left(D_p'\alpha_p' + D_p''\alpha_p'' + \cdots D_p^{(q)}\alpha_p^{(q)}\right) \neq \bar{0}, \tag{112}$$

and this means also that the relation

$$N^{p-1}\left(D_p'\alpha_p' + D_p''\alpha_p'' + \cdots D_p^{(q)}\alpha_p^{(q)}\right) = \bar{0}, \tag{113}$$

necessarily implies $\alpha_p' = \alpha_p'' = \cdots = \alpha_p^{(q)} = 0$ To each element $D_p^{(i)}$, one can now also construct the associated elements

$D^{(i)}_{p-1}$, $D^{(i)}_{p-2}$, ... $D^{(i)}_2$, $D^{(i)}_1$ $(\neq \bar{0})$ according to the scheme (105), or

$$D^{(i)}_j = N D^{(i)}_{j+1} = \cdots = N^{p-j} D^{(i)}_p . \tag{114}$$

It can now easily be shown that the pq elements $D^{(i)}_j$, for i = 1, 2, .. q and j = 1, 2, ... p, form a linearly independent set. For the proof, we will consider the linear relation

$$\sum_{i=1}^{q} \sum_{j=1}^{p} D^{(i)}_j \alpha^{(i)}_j = \bar{0} . \tag{115}$$

Multiplication to the left by N^{p-1} will annihilate all terms except those for j = p and gives

$$N^{p-1} \left(D'_p \alpha'_p + D''_p \alpha''_p + \cdots D^{(q)}_p \alpha^{(q)}_p \right) = \bar{0} . \tag{116}$$

According to (113), one has then $\alpha'_p = \alpha''_p = \cdots = \alpha^{(q)}_p = 0$, and relation (115) may be simplified to the form

$$\sum_{i=1}^{q} \sum_{j=1}^{p-1} D^{(i)}_j \alpha^{(i)}_j = \bar{0} . \tag{117}$$

Multiplication to the left by N^{p-2} will annihilate all terms except those for j = p-1 and, using (114), one obtains

$$N^{p-1} \left(D'_p \alpha'_{p-1} + D''_p \alpha''_{p-1} + \cdots D^{(q)}_p \alpha^{(q)}_{p-1} \right) = \bar{0} \tag{118}$$

i.e. $\alpha'_{p-1} = \alpha''_{p-1} = \cdots \alpha^{(q)}_{p-1} = 0$. Proceeding in this way, one finds finally that all the coefficients $\alpha^{(i)}_j$ are necessarily vanishing, which proves the theorem.

Since the number of independent elements cannot exceed the order of the space, one has the condition $pq \leq n$. If it happens that $pq = n$, one chooses the elements $D_j^{(i)}$ as a basis, and, according to (114) or the corresponding relation

$$N D_j^{(i)} = D_{j-1}^{(i)}, \qquad \text{for } j = 1, 2, \ldots p \tag{119}$$

one finds that N has a matrix representation which consists of q Jordan blocks of order p.

If $pq < n$, we will consider the subspace V_{p-1} of the total space which is such that, for every element A in V_{p-1}, one has $N^{p-2}A \neq 0$. Let the order of this subspace be r. One has already q independent elements $D'_{p-1}, D''_{p-1}, \ldots D^{(q)}_{p-1}$, belonging to this space, which means that $r \geq q$. If $r = q$, one proceeds to consider the space V_{p-2}. If, on the other hand $r > q$, one selects $r' = (r-q)$ elements $E'_{p-1}, E''_{p-1}, \ldots E^{(r-q)}_{p-1}$, such that together with the elements $D'_{p-1}, D''_{p-1}, \ldots D^{(q)}_{p-1}$ they form a linearly independent set which spans the subspace V_{p-1}. Introducing the elements $E^{(k)}_{p-2}, E^{(k)}_{p-3}, \ldots E^{(k)}_2, E^{(k)}_1$ through the relations

$$E_l^{(k)} = N E_{l+1}^{(k)} = \cdots = N^{p-1-l} E_{p-1}^{(k)}, \qquad k = 1, 2 \ldots r', \; l = 1, 2 \ldots p-1 \tag{120}$$

one obtains a set of $pq + (p-1)r'$ elements $D_j^{(i)}, E^{(k)}$ which are easily shown to be linearly independent. If $n = pq + (p-1)r'$, one can choose this set as a basis and obtains a matrix representation of N which consists of q Jordan blocks of order p, and r' Jordan blocks of order $(p-1)$. On the other hand, if $n > pq + (p-1)r'$, one proceeds by considering the subspace V_{p-2}, etc. In this way, one proves that there exists a special basis in which every nilpotent matrix N of order p has a representation which consists of a diagonal series of Jordan blocks characterized by their orders or Segré characteristics.

In conclusion, we observe that a nilpotent matrix N has only the eigenvalue 0 which has the multiplicity n, and that the associated Cayley-Hamilton equation is hence

$$N^n = 0 . \tag{121}$$

However, if the largest Jordan block has the order $m \leq n$, the nil-potent matrix satisfies actually also the relation

$$N^{m} = 0 , \tag{122}$$

which for $m < n$ has a lower degree than (121), and one says that (122) is a "reduced" Cayley-Hamilton equation. This concept will be of fundamental importance in the following.

Classical canonical form of a matrix in the general case. - In the two previous sections, we have particularly considered the case of a linear operator T which has only distinct eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_m$, and we have shown that the matrix $\mathbf{T}_x$ in an arbitrary representation $\mathbf{X}$ may be brought to diagonal form by a suitable similarity transformation (79).

In this section, we will consider the general case in which one or more eigenvalues may be degenerate corresponding to multiple roots to the characteristic polynomial (70). In such a case, it is usually not possible to bring the matrix to diagonal form, but other simplifications may instead be carried out by feasible similarity transformations.

In order to study the effect of a degeneracy, we will first consider the case of a single eigenvalue λ having the multiplicity n. According to the general Cayley-Hamilton theorem (80), the operator T satisfies the algebraic equation

$$(T - \lambda \cdot I)^{m} = 0 . \tag{123}$$

This implies that the operator $N = T - \lambda \cdot I$ is a nilpotent operator of an order $p \leq n$. Since the operator N may be represented by a set of Jordan blocks, there exists apparently a basis in which the operator T may be represented by a set of blocks having the eigenvalue λ in the diagonal and the number 1 in the diagonal above.[10] According to (111), we obtain for $n = 4$ the simple examples:

$$\underset{(4)}{\overset{p=4}{\left[\begin{array}{cccc}\lambda&1&&\\&\lambda&1&\\&&\lambda&1\\&&&\lambda\end{array}\right]}}\quad \underset{(3,1)}{\overset{p=3}{\left[\begin{array}{ccc|c}\lambda&1&&\\&\lambda&1&\\&&\lambda&\\\hline&&&\lambda\end{array}\right]}}\quad \underset{(2,2)}{\overset{p=2}{\left[\begin{array}{cc|cc}\lambda&1&&\\&\lambda&&\\\hline&&\lambda&1\\&&&\lambda\end{array}\right]}}\quad \underset{(2,1,1)}{\overset{p=2}{\left[\begin{array}{cc|c|c}\lambda&1&&\\&\lambda&&\\\hline&&\lambda&\\\hline&&&\lambda\end{array}\right]}}\quad \underset{(1,1,1,1)}{\overset{p=1}{\left[\begin{array}{c|c|c|c}\lambda&&&\\\hline&\lambda&&\\\hline&&\lambda&\\\hline&&&\lambda\end{array}\right]}} \tag{124}$$

which may be sufficient as an illustration. Again we note that the order p is determined by the number m, which is the order of the largest Jordan block, i.e. the largest Segré characteristic, so that

$$(T-\lambda\cdot I)^{p} = 0 . \tag{125}$$

For $p < n$, one obtains hence a reduced Cayley-Hamilton equation.

Next, we will consider the general case when the linear operator T has eigenvalues $\lambda_1, \lambda_2, \lambda_3, \dots$ of the multiplicity $g_1, g_2, g_3, \dots$, respectively, with $g_1 + g_2 + g_3 + \dots = n$. In this case, the Cayley-Hamilton theorem (80) may be written in the form

$$(T-\lambda_1\cdot I)^{g_1}(T-\lambda_2\cdot I)^{g_2}\dots(T-\lambda_k I)^{g_k} = 0 . \tag{126}$$

The characteristic polynomial in the complex variable z may be written in the form

$$P(z) \equiv \prod_j (\lambda_j - z)^{g_j} , \tag{127}$$

and, using the technique for developing into partial fractions, one obtains

$$\frac{1}{P(z)} \equiv \sum_k \frac{p_k(z)}{(\lambda_k - z)^{g_k}} , \tag{128}$$

where $p_k(z)$ is a polynomial of degree less than g_k. This gives directly the identity

$$1 \equiv \sum_k \frac{p_k(z)}{(\lambda_k - z)^{g_k}} P(z) =$$

$$\equiv \sum_k p_k(z) \prod_{j \neq k}' (\lambda_j - z)^{g_j} , \tag{129}$$

which is valid even if one substitutes the linear operator T instead of z and the identity operator I instead of 1:

$$I = \sum_k p_k(T) \prod_{j \neq k}' (\lambda_j \cdot I - T)^{g_j} . \tag{130}$$

This implies that one can subdivide the original space V of order n into subspaces $W_1, W_2, \ldots W_k$ associated with the individual eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_k$ by the formulas

$$V = \sum_k W_k , \tag{131}$$

$$W_k = p_k(T) \prod_{j \neq k}' (\lambda_j \cdot I - T)^{g_j} \cdot V$$

i.e. every element A of V can be uniquely decomposed in this way. By using the Cayley-Hamilton theorem (126), one gets immediately

$$(T - \lambda_k \cdot I)^{g_k} W_k = \bar{0} , \tag{132}$$

i.e. the operator $N_k = T - \lambda_k I$ is nilpotent of an order not exceeding g_k within the subspace W_k. By choosing a convenient basis within W_k, one can now represent the operator N_k in the classical canonical form previously discussed. The order of the subspace W_k must be exactly equal to the multiplicity g_k. If the order would be higher, one could construct a secular determinant for T in which the eigenvalue λ_k would have a higher multiplicity than g_k, which would be a contradiction. On the other hand, if the order would be lower than g_k, the sum of the orders of the subspaces W_k would be lower than n, which is another contradiction.

By using the operators in (129), it is easily shown that elements associated with different subspaces W_k are linearly independent. The bases used to span the subspaces W_k may hence be put together to form a basis for the complete space V. In this basis, the matrix for T will hence consist of a series of diagonal blocks of the type (124). Each block is conveniently characterized by the eigenvalue λ_k and the associated Segrê characteristics describing the form of the diagonal immediately above the main diagonal.

Reduced Cayley-Hamilton equation. - Let us consider an eigenvalue λ_k having the degeneracy g_k and the largest Segrê characteristic m_k. The Cayley-Hamilton theorem has the product form

$$\prod_j (\lambda_j \cdot I - T)^{g_j} = 0. \tag{133}$$

However, since the largest Jordan block associated with the eigenvalue λ_k has the order m_k, it is directly seen that the matrix $\mathbf{T}$, and hence also the operator T, satisfies a reduced Cayley-Hamilton equation of the form

$$\boxed{\prod_j (\lambda_j \cdot I - T)^{m_j} = 0\,.} \tag{134}$$

The associated polynomial $F(z) = \prod_j (\lambda_j - z)^{m_j}$ is often called the minimal-polynomial associated with the operator T.

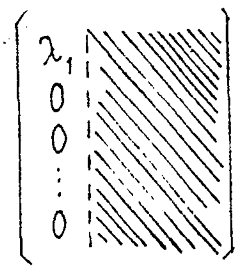
This concept is of particular importance in treating infinite linear spaces. Even if the degeneracies g_k become infinite and the Cayley-Hamilton theorem (133) loses its meaning, it may happen that the numbers m_j stay finite and that the reduced equation (134) exists. We will return to this problem in a later section.

Triangularization of a matrix. - Since the problem of the simplest possible form of a matrix representation for a linear operator T is of fundamental importance, we will here briefly reconsider it from another point of view.

Let us again start from the eigenvalue problem, $TC = \lambda C$. The theory of systems of linear equations tells us that, for each root λ_k to the secular equation (67), there exists at least one eigenelement C_k. Using the same technique as in (74), one can easily show that eigenelements associated with different eigenvalues are necessarily linearly independent.

Starting from the eigenvalue λ_1 and the associated eigenelement, we will now choose a set of linearly independent elements $C_1, Y_2, Y_3, \ldots Y_n$ as a basis. The operator T gets then a matrix representation of the type

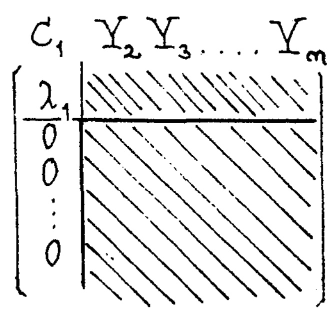
$\mathbf{T}$ =



(135)

where the form of the first column depends on the relation $TC_1 = C_1\lambda_1$ and its interpretation in matrix form according to (31). This implies that, by a similarity transformation, one can bring any quadratic matrix $\mathbf{T}_x$ to the special form (135) with only zeros in the first column below the diagonal. Let us now partition the matrix (135) and consider particularly the quadratic matrix of order (n-1) associated with the elements $Y_2, Y_3, \ldots Y_n$:

$\mathbf{T}$ =



(135')

By a similarity transformation, this matrix may now be brought to the form (135) and, repeating the procedure, we are finally led to the matrix:

$$\begin{bmatrix} \lambda_1 & & & & \\ 0 & \lambda_1' & & & \\ 0 & 0 & \lambda_1'' & & \\ 0 & 0 & 0 & \lambda_1''' & \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \tag{136}$$

which is characterized by the fact that it has only zero's below the entire diagonal. Since the associated secular determinant takes the form $(\lambda_1 - z)(\lambda_1' - z)(\lambda_1'' - z)\ldots = 0$ and has the roots $\lambda_1, \lambda_1', \lambda_1'', \ldots$, one can conclude that these numbers must be equal to the original eigenvalues.

This simple procedure is called a <u>triangularization</u> of a matrix, and it shows that any matrix may be brought to a triangular form of the type (136) with the eigenvalues in the main diagonal and only zeros below it by means of a suitable similarity transformation.

Let us now consider a degenerate eigenvalue λ_1 and arrange the triangularization, so that this eigenvalue is repeated consecutively along the diagonal as many times as its multiplicity. The corresponding basic elements will be denoted by C_1', C_1'', C_1''', ... , and the matrix representation (136) takes the form

$$\mathbf{T} = \begin{array}{c} \begin{array}{ccccc} C_1' & C_1'' & C_1''' & C_1'''' & \ldots \end{array} \\ \begin{bmatrix} \lambda_1 & \alpha_{12} & \alpha_{13} & \alpha_{14} & \ldots \\ 0 & \lambda_1 & \alpha_{23} & \alpha_{24} & \ldots \\ 0 & 0 & \lambda_1 & \alpha_{34} & \ldots \\ 0 & 0 & 0 & \lambda_1 & \ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \end{array} , \tag{137}$$

where the elements α_{kl} are not yet determined and may be vanishing or non-vanishing depending on the character of the operator T.

Let us first consider the case that all the elements α_{kl} are non-vanishing. According to (31), one has

$$\begin{cases} TC_1' = C_1' \lambda_1 , \\ TC_1'' = C_1' \alpha_{12} + C_1'' \lambda_1 , \\ TC_1''' = C_1' \alpha_{13} + C_1'' \alpha_{23} + C_1''' \lambda_1 , \\ \ldots\ldots\ldots \end{cases}$$

or

$$
\begin{aligned}
(T - \lambda_1 \cdot I)\, C_1' &= \bar{0} \,, \\
(T - \lambda_1 \cdot I)\, C_1'' &= C_1' \alpha_{12} \,, \\
(T - \lambda_1 \cdot I)\, C_1''' &= C_1' \alpha_{13} + C_1'' \alpha_{23} \,, \\
&\ldots\ldots\ldots
\end{aligned}
\tag{138}
$$

This gives further the relations

$$(T - \lambda_1 \cdot I)\, C_1' = \bar{0}\,, \quad (T - \lambda_1 \cdot I)^2 C_1'' = \bar{0}\,, \quad (T - \lambda_1 \cdot I)^3 C_1''' = \bar{0}\,, \ldots \tag{139}$$

showing that, if the eigenelement C_1' is annihilated by the operator $(T - \lambda_1 \cdot I)$, the higher basic elements C_1'', C_1''', ... are apparently annihilated by the <u>powers</u> of this operator. Within the subspace spanned by C_1', C_1'', C_1''', ... the operator $(T - \lambda_1 \cdot I)$ is hence nilpotent of an order which does not exceed the multiplicity g_1.

The next step is to consider the case that not all the elements α_{kl} are non-vanishing, and one is led in this way to the concepts of block formation, the Segrê characteristics, and the reduced Cayley-Hamilton equation. The question of the transformation to the "classical canonical form" is treated in an excellent way in many textbooks[11], in which one may find the pertinent literature references and further details.

6. COMPONENT ANALYSIS IN A LINEAR SPACE OF INFINITE ORDER

In the previous sections, we have explicitly confined our interest to linear spaces of finite order n, and all the conclusions have been based on the concepts of linear independence and the existence of a basis. In order to generalize these considerations to linear spaces of an infinite order, one has to deal with complicated convergence problems which are the subject of e.g. the theory of Hilbert space. In this situation, it seems rather remarkable that there still exists a series of theorems about infinite spaces which are non-trivial and of fundamental importance in quantum theory. This depends on the fact that, even if the operators involved have an infinite number of eigenvalues, these are situated only in a <u>finite</u> number of points in the complex plane each of which may have an infinite degeneracy. To illustrate the problem, we will start with a simple example.

<u>Exchange operator P_{12}.</u> - Let us consider the linear space formed by all functions $\Phi = \Phi(1,2)$ of two coordinates; x_1 and x_2. Such a space cannot be spanned by a finite number of elements, and it has hence an infinite order. We will further consider a linear operator $P = P_{12}$ which interchanges the two coordinates, so that

$$P\,\Phi(1,2) = \Phi(2,1) \tag{140}$$

This is a permutation operator identical with the simplest "exchange operator" in quantum theory. Using (23), one finds that two interchanges give back the original element, i.e.

$$P^2 = I \tag{141}$$

The eigenvalue problem has the form $PC = \lambda C$, and one obtains

$P^2 C = \lambda^2 C = C$, showing that $\lambda^2 = 1$. The eigenvalues are hence $\lambda = \pm 1$.

In order to proceed, we will use an identity which is easily found by inspection:

$$\Phi(1,2) \equiv \frac{1}{2}\left[\Phi(1,2) + \Phi(2,1)\right] + \frac{1}{2}\left[\Phi(1,2) - \Phi(2,1)\right] \tag{142}$$

It appears that the first term in the right-hand member is an eigenelement to P associated with the eigenvalue +1, whereas the second term is an eigenelement associated with the eigenvalue -1. There are apparently only two eigenvalues, but both of them are infinitely degenerate.

The symmetric and antisymmetric element in (142) may be obtained from the original element by means of the operators:

$$\mathcal{O}_{+1} = \frac{1}{2}(I + P) \qquad \mathcal{O}_{-1} = \frac{1}{2}(I - P) \tag{143}$$

and, by using (141), it is easily shown that they satisfy the algebraic identities:

$$P\mathcal{O}_{+1} = \mathcal{O}_{+1}, \qquad P\mathcal{O}_{-1} = -\mathcal{O}_{-1}; \tag{144}$$

$$\mathcal{O}_{+1}^2 = \mathcal{O}_{+1}, \qquad \mathcal{O}_{-1}^2 = \mathcal{O}_{-1}, \qquad \mathcal{O}_{+1}\mathcal{O}_{-1} = 0; \tag{145}$$

$$I = \mathcal{O}_{+1} + \mathcal{O}_{-1} \tag{146}$$

i.e. they are mutually exclusive projection operators, which are eigenoperators and form a resolution of the identity. By means of these operators, one can split the entire space V into two subspaces $O_{+1}V$ and $O_{-1}V$, each of an infinite order, which are directly associated with the eigenvalues $\lambda = +1$ and $\lambda = -1$, respectively. The relations (144)-(146) are completely analogous to the relations (91)-(93) and represent some form of generalization to a space of an infinite order. In the following, we will try to systematize this approach.

<u>Projection operators and resolution of the identity based on the use of the reduced Cayley-Hamilton equation.</u> - Let us consider an infinite linear space and a linear operator T such that it has all its eigenvalues situated in a <u>finite</u> number of points $\lambda_1, \lambda_2, \dots \lambda_m$ in the complex plane. Each one of these eigenvalues λ_k may hence be infinitely degenerate ($g_k = \infty$), but we will assume that the largest Segré characteristic m_k is always <u>finite</u>, and we will start by considering the case $m_1 = m_2 = \dots = m_n = 1$. This implies that the operator T satisfies a reduced Cayley-Hamilton equation of the type (134), in which each eigenvalue factor occurs only once. The associated minimal polynomial is hence

$$F(z) = \prod_{j=1}^{m} (\lambda_j - z) , \tag{147}$$

and the basic assumption may be written in the form

$$\boxed{F(T) = \prod_{j=1}^{m} (\lambda_j \cdot I - T) = 0} \tag{148}$$

Our treatment will be based solely on this operator relation. Equation (141) is of this type, and we will later see that many other fundamental operators in quantum theory fulfil similar relations.

In analogy with (103), we will now define a set of operators $O_1, O_2, \dots O_n$ by means of the product formula

$$O_k = \prod_{j \neq k} \frac{T - \lambda_j \cdot I}{\lambda_k - \lambda_j} \tag{149}$$

Since O_k consists of all the factors which occur in F(T) except for the single factor ($T - \lambda_k \cdot I$), one obtains immediately $(T - \lambda_k \cdot I) O_k = 0$, or

$$T O_k = \lambda_k O_k , \tag{150}$$

showing that O_k is an eigenoperator to T. One may also write (149) in the form,

$$\mathrm{O}_k = \prod_{j \neq k} \left\{ \mathrm{I} + \frac{T - \lambda_k \cdot \mathrm{I}}{\lambda_k - \lambda_j} \right\} , \tag{151}$$

and, using $(T - \lambda_k \cdot \mathrm{I})\,\mathrm{O}_k = \bar{0}$, one finds directly

$$\mathrm{O}_k^2 = \mathrm{O}_k \; ; \qquad \mathrm{O}_k \mathrm{O}_l = 0 , \quad k \neq l . \tag{152}$$

The operators O_k are hence idempotent and mutually exclusive.

It is now easily shown that the projection operators O_k defined by (149) also form a "resolution of the identity". Since one has no expansion theorem to rely on, it is necessary to proceed in a completely different way. In addition to the minimal polynomial $F(z)$, we will now consider also the polynomials $O_k(z)$ defined by the relations

$$\mathrm{O}_k(z) = \frac{F(z)}{(z - \lambda_k) F'(\lambda_k)} = \prod_{j \neq k} \frac{z - \lambda_j}{\lambda_k - \lambda_j} . \tag{153}$$

These are polynomials of degree (n-1) which have the value 1 for $z = \lambda_k$ and the value 0 for $z = \lambda_j \; (j \neq k)$, and they are thus Legendre "interpolation polynomials". Let us further consider the auxiliary function

$$G(z) \equiv 1 - \sum_{k=1}^{m} \mathrm{O}_k(z) . \tag{154}$$

Since this is a polynomial of degree (n-1) having the value zero in the n points $z = \lambda_1, \lambda_2, \ldots \lambda_m$, one obtains $G(z) \equiv 0$. This identity is valid in terms of the complex variable z, but it remains valid even if one replaces z by the operator T and the number 1 by the identity operator I. Hence one has

$$\mathrm{I} \equiv \sum_{k=1}^{m} \mathrm{O}_k(T) , \tag{155}$$

which is the "resolution of the identity" desired.

Let us now investigate how the operator relations (150), (152), and (155) may be utilized for a treatment of the infinite linear space. If A is an arbitrary element of the space, one obtains by using (155) the following decomposition of the element:

$$A = I \cdot A = \left(\sum_{k=1}^{m} O_k \right) A = \sum_{k=1}^{m} O_k A = \sum_{k=1}^{m} A_k , \tag{156}$$

where $A_k = O_k A$. Using (150), one gets

$$T A_k = \lambda_k A_k , \tag{157}$$

which shows that (156) is a resolution of A into eigenelements of T. According to (107), one has further $O_k A_k = A_k$, whereas $O_k A_\ell = \overline{0}$ for $k \neq \ell$. Using this property, one can easily show that the decomposition of A into eigenelements is unique, for, if there would be two relations $A = \sum_k A'_k = \sum_k A''_k$ having components satisfying (157), multiplication by O_ℓ would give $A'_\ell = A''_\ell$.

Even if the resolution (156) contains a sum, it should not be confused with an expansion theorem of the type (5) or (76) derived by means of the concept of a basis. Instead, it is more appropriate to describe (156) as a <u>component analysis</u> of an element A in terms of eigenelements to T, and the component $A_k = O_k A$ is said to be the projection of the element A on the eigenspace of T associated with the eigenvalue λ_k .

By means of the projection operators $O_1, O_2, \ldots O_n$, it is further possible to split the given space V into subspaces $V_1, V_2, \ldots V_n$ associated with the various eigenvalues:

$$V = \sum_{k=1}^{m} V_k , \qquad V_k = O_k V , \qquad T V_k = \lambda_k V_k , \tag{158}$$

and we will describe this procedure as a "splitting of V after eigenvalues to T".

In conclusion, we note that, applying T to (155) and using (150), one obtains

$$T = \sum_{k=1}^{m} \lambda_k O_k , \tag{159}$$

which is a "spectral resolution" of the operator T corresponding to (94). If f(z) is an arbitrary polynomial in the complex variable z, one gets further

$$f(T) = \sum_{k=1}^{m} f(\lambda_k)\, O_k , \tag{160}$$

and, from the polynomial, one can then proceed to consider algebraic functions.

As an illustration of the projection technique, we will now derive a simple theorem. Let A be an arbitrary element of the infinite space, and let us consider the linear manifold spanned by the elements

$$A,\ TA,\ T^2 A, \ldots \quad T^{m-1} A \tag{161}$$

Taking the projection O_k of an arbitrary element out of this manifold and using the relation $TO_k = O_k T = \lambda_k O_k$, one obtains

$$\begin{aligned} & O_k \left(a_0 A + a_1 TA + a_2 T^2 A + \cdots + a_{m-1} T^{m-1} A \right) = \\ = \; & O_k \left(a_0 A + a_1 \lambda_k A + a_2 \lambda_k^2 A + \cdots + a_{m-1} \lambda_k^{m-1} A \right) = \\ = \; & \left(a_0 + a_1 \lambda_k + a_2 \lambda_k^2 + \cdots + a_{m-1} \lambda_k^{m-1} \right) O_k A \end{aligned} \tag{162}$$

Hence the projection of an arbitrary element of the manifold is proportional to the projection of the element A itself. In quantum theory, this theorem is often quite useful in different connections.

The minimal polynomial (147) is a special case of the integer functions, and it is an interesting problem to investigate whether our approach may be generalized also to the case of an infinite number of eigenvalues λ_k by using the theory of infinite products with and without converging factors.

Cyclic operators. - As an example of the method described, we will consider the eigenvalue problem of the cyclic operators which are characterized by the relation

$$T^G = I \, , \tag{163}$$

where G is an integer. From the relation $TC = \lambda C$ follows directly $T^G C = \lambda^G C = C$ or $\lambda^G = 1$, which gives the eigenvalues

$$\lambda_k = e^{2\pi i k/G} \tag{164}$$

for k = 0, 1, 2, ... G-1. For the interpolation polynomials (153), one obtains

$$\begin{aligned} O_k(z)' &= \frac{1 - z^G}{(\lambda_k - z) G \lambda_k^{G-1}} = \frac{1}{G} \, \frac{1 - \lambda_k^{-G} z^G}{1 - \lambda_k^{-1} z} = \\ &= \frac{1}{G} \sum_{l=0}^{G-1} \lambda_k^{-l} z^l \, . \end{aligned} \tag{165}$$

For the eigenprojectors, this gives

$$\boxed{O_k = \frac{1}{G} \sum_{l=0}^{G-1} e^{-2\pi i k l/G} \, T^l} \tag{166}$$

By means of these operators, it is now possible to split the infinite linear space V into G subspaces $V_0, V_1, \dots V_{G-1}$ of infinite order associated with the eigenvalues $\lambda_0, \lambda_1, \dots \lambda_{G-1}$, respectively.

The eigenvalue relation (150) gives directly

$$T \, O_k = e^{2\pi i k/G} \, O_k \, , \tag{167}$$

which is equivalent to the "Bloch condition" in quantum theory. According to (152) and (155), one has further

$$\mathcal{O}_k^2 = \mathcal{O}_k, \qquad \mathcal{O}_k\mathcal{O}_l = 0, \quad k \neq l \tag{168}$$

$$I = \sum_k \mathcal{O}_k, \tag{169}$$

which relations may now be checked explicitly. For $G = 2$, one obtains the special case of the exchange operator $P = P_{12}$ defined by (140) and (141).

Translations. - Let us start by considering a linear space consisting of all functions $\Phi(x)$ of a single variable x, and let T be a translational operator connected with the length a defined by the relation:

$$T\,\Phi(x) = \Phi(x+a). \tag{170}$$

In order to proceed, we will assume that all the functions Φ under consideration fulfil the Born- von Kârmân boundary condition:

$$\Phi(x+Ga) = \Phi(x), \tag{171}$$

where G is a very large integer. Using (170), one can write this condition in the form $(T^G - I)\,\Phi \equiv \bar{0}$. which means that T is a cyclic operator of order G for all functions satisfying the periodicity condition. The eigenvalues and eigenprojectors are hence given by the relations (164) and (166), respectively. By using the projection technique, an arbitrary element Φ may now be resolved into eigenfunctions to T, so that

$$\Phi = \sum_{\varkappa=0}^{G-1} \Phi_\varkappa, \tag{172}$$

$$\Phi_\varkappa = \mathcal{O}_\varkappa \Phi = \frac{1}{G} \sum_{\nu=0}^{G-1} e^{-2\pi i \nu\varkappa/G}\, \Phi(x+\nu a). \tag{173}$$

The components $\Phi_\varkappa$ are identical with the well-known Bloch functions.[12]

<u>Projection splitting in the case of a general reduced Cayley-Hamilton equation.</u> - Again we will consider an infinite linear space V and an operator T having all its eigenvalues situated in a finite number of points $\lambda_1, \lambda_2, \ldots \lambda_m$. Let us consider a more general case than before and assume that the minimal polynomial has the form

$$F(z) \equiv \prod_{j=1}^{m} (\lambda_j - z)^{m_j} , \tag{148}$$

and that T satisfies the reduced Cayley-Hamilton equation

$$F(T) = \prod_{j=1}^{m} (\lambda_j \cdot I - T)^{m_j} = 0 . \tag{149}$$

In order to derive a "resolution of the identity", we will now, in analogy to (128), study the algebraic identity

$$\frac{1}{F(z)} \equiv \sum_{k=1}^{m} \frac{q_j(z)}{(\lambda_j - z)^{m_j}} , \tag{150}$$

or

$$1 \equiv \sum_{k=1}^{m} q_j(z) \frac{F(z)}{(\lambda_j - z)^{m_j}} , \tag{151}$$

where $q_j(z)$ is a polynomial of degree less than m_j. Introducing the operator

$$\mathbb{O}_k(T) = q_k(T) \prod_{j \neq k} (\lambda_j \cdot I - T)^{m_j} , \tag{152}$$

one gets immediately, according to (149) and (151), the relations

$$(T - \lambda_k \cdot I)^{m_k} \mathbb{O}_k = 0 \tag{153}$$

$$I = \sum_k \mathbb{O}_k \tag{154}$$

Since further $O_k O_\ell = 0$ for $k \neq \ell$ according to (149), one obtains

$$O_k = O_k \sum_\ell O_\ell - \sum_\ell O_k O_\ell = O_k^2 , \qquad (155)$$

showing that O_k is idempotent anu hence a projection operator. This leads to a unique component-analysis

$$A = \sum_k A_k , \qquad A_k = O_k A , \qquad (156)$$

$$(T - \lambda_k \cdot I)^{m_k} A_k = 0 \qquad (157)$$

Introducing the subspace $V_k = O_k V$, we have thus found that the operator $(T - \lambda_k \cdot I)$ is nilpotent of order m_k with respect to this subspace. For $m_1 = m_2 = \ldots = m_n = 1$, we obtain the formulas previously derived.

7. SIMULTANEOUS SPLITTING OF FINITE AND INFINITE LINEAR SPACES WITH RESPECT TO A SET OF COMMUTING LINEAR OPERATORS

Let us start from the concept of "stability": a linear subspace W is said to be <u>stable</u> under the operation T, if the subspace TW belongs entirely to W. Since $TC = \lambda C$, an eigenelement C represents always a stable subspace of the first order.

Let further R be another linear operator which <u>commutes</u> with T, so that

$$RT = TR \tag{158}$$

It is now easily seen that, if V_k is an eigenspace to T associated with the eigenvalue λ_k, then RV_k belongs also to V_k, i.e. the eigenspace V_k is stable under the operation R. Since $TV_k = \lambda_k V_k$, one has

$$\begin{aligned} T(RV_k) &= (TR)V_k = (RT)V_k = R(TV_k) = \\ &= R(\lambda_k V_k) = \lambda_k (RV_k), \end{aligned} \tag{159}$$

which completes the proof. However, since the subspace V_k is stable under the operation R, it is now possible to consider the eigenvalue problem of R within this subspace. The procedure will lead to eigenelements C_{kl}, which are <u>simultaneous eigenelements</u> to the operators T and R:

$$TC_{kl} = \lambda_k C_{kl}, \qquad RC_{kl} = \mu_l C_{kl}. \tag{160}$$

The circumstances will be particularly simple, if both T and R have only a finite number of eigenvalues, which may be even infinitely degenerate but all have the largest Segré characteristic $m = 1$, so that they satisfy reduced Cayley-Hamilton equations of the type (148). According to (149), one may then introduce the projection operators associated with T and R, respectively:

$$\mathbb{O}_k(T) = \prod_{j \neq k} \frac{T - \lambda_j \cdot I}{\lambda_k - \lambda_j} , \tag{161}$$

$$\mathbb{O}_l(R) = \prod_{i \neq l} \frac{R - \mu_i \cdot I}{\mu_l - \mu_i} , \tag{162}$$

which both satisfy resolutions of the identity (155). One obtains directly

$$\begin{aligned} I &= \{ \sum_k \mathbb{O}_k(T) \} \{ \sum_l \mathbb{O}_l(R) \} = \\ &= \sum_k \sum_l \mathbb{O}_k(T) \mathbb{O}_l(R) , \end{aligned} \tag{163}$$

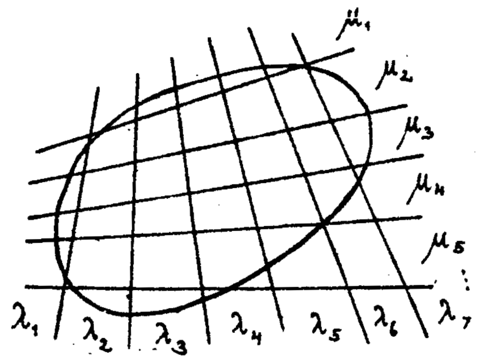
where the terms in the double sum

$$\mathbb{O}_{kl} = \mathbb{O}_k(T) \mathbb{O}_l(R) \tag{164}$$

are again projection operators which are idempotent, mutually exclusive, and satisfy the relations

$$T \mathbb{O}_{kl} = \lambda_k \mathbb{O}_{kl} , \quad R \mathbb{O}_{kl} = \mu_l \mathbb{O}_{kl} , \tag{165}$$

according to (158). This implies that the operators O_{kl} are simultaneous eigenoperators to T and R.



Through the double sum in (163), one obtains a simultaneous splitting of the space V into subspaces V_{kl} which are simultaneous eigenspaces to the operators T and R:

$$V = \sum_{k,l} V_{kl} = \sum_{k,l} \mathbb{O}_{kl} V \tag{166}$$

Some of these spaces may be empty, i.e. contain only the element $\overline{0}$, but, for an infinite space V, the subspaces $V_{k\ell}$ are usually of an infinite order.

One can generalize this idea still further for, if there exists a set of linear operators T, R, S, ... which are mutually commuting, one can carry out a splitting of V into simultaneous eigenspaces to these operators by means of projection operators of the type

$$\mathcal{O}_{k\ell m\ldots} = \mathcal{O}_k(T)\,\mathcal{O}_\ell(R)\,\mathcal{O}_m(S)\ldots, \tag{167}$$

which form a resolution of the identity.

Case of $m \neq 1$. - Let us also consider the case when T and R are linear operators which do not necessarily have $m = 1$. The reduced Cayley-Hamilton equation takes now the form (149). Introducing the projection operators $O_k(T)$ and $O_\ell(R)$ according to (152), one obtains again a resolution of the identity of the type (163), where the product operators $O_{k\ell} = O_k(T)\,O_\ell(R)$ are now satisfying the relations

$$(T - \lambda_k \cdot I)^{m_k}\,(R - \mu_\ell \cdot I)^{m_\ell}\,\mathcal{O}_{k\ell} = 0. \tag{168}$$

Even in this case, it is hence possible to obtain a splitting of V into eigenspaces $V_{k\ell} = O_{k\ell}V$ which are associated with the pair (λ_k, μ_ℓ) of characteristic numbers.

Translations in three dimensions. - As an application of this splitting technique, we will consider the linear space of all functions $\Phi = \Phi(\mathbf{r})$ of a three-dimensional variable $\mathbf{r}$ subject to the three fundamental translations T_1, T_2, and T_3 defined by:

$$\begin{aligned} T_1\,\Phi(\mathbf{r}) &= \Phi(\mathbf{r} + \mathbf{a}_1), \\ T_2\,\Phi(\mathbf{r}) &= \Phi(\mathbf{r} + \mathbf{a}_2), \\ T_3\,\Phi(\mathbf{r}) &= \Phi(\mathbf{r} + \mathbf{a}_3). \end{aligned} \tag{169}$$

Here $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are the primitive translations, which form a parallellepiped of volume $V_a = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$. It is easily shown that the three operators T_1, T_2, T_3 mutually commute. The vector $\mathbf{m} = \mu_1 \mathbf{a}_1 + \mu_2 \mathbf{a}_2 + \mu_3 \mathbf{a}_3$ is called a general translation and, for the associated operator $T(\mathbf{m})$, one has the connection formula $T(\mathbf{m}) = T_1^{\mu_1} T_2^{\mu_2} T_3^{\mu_3}$. The treatment is simplified by the assumption, that all functions Φ under consideration satisfy the Born-von Kármán boundary condition:

$$\Phi(\mathbf{r} + G_\nu \mathbf{a}_\nu) = \Phi(\mathbf{r}), \qquad \nu = 1,2,3 \tag{170}$$

where (G_1, G_2, G_3) is a triple of large integers. This leads to the conditions

$$T_1^{G_1} \equiv I, \quad T_2^{G_2} \equiv I, \quad T_3^{G_3} \equiv I, \tag{171}$$

and the eigenvalues and the associated projection operators are hence given by (164) and (166), respectively. Introducing the simultaneous eigenoperators to T_1, T_2, T_3 according to (167), we obtain

$$\begin{aligned} O(\varkappa_1, \varkappa_2, \varkappa_3) &= O_{\varkappa_1}(T_1)\, O_{\varkappa_2}(T_2)\, O_{\varkappa_3}(T_3) = \\ &= \prod_{\nu=1}^{3} \left\{ G_\nu^{-1} \sum_{\mu_\nu=0}^{G_\nu - 1} e^{-2\pi i \varkappa_\nu \mu_\nu / G_\nu}\, T_\nu^{\mu_\nu} \right\} = \\ &= (G_1 G_2 G_3)^{-1} \sum_{\mu_1,\mu_2,\mu_3=0}^{G_\nu - 1} e^{-2\pi i \left(\frac{\varkappa_1 \mu_1}{G_1} + \frac{\varkappa_2 \mu_2}{G_2} + \frac{\varkappa_3 \mu_3}{G_3} \right)} T_1^{\mu_1} T_2^{\mu_2} T_3^{\mu_3}, \end{aligned} \tag{172}$$

where $(\varkappa_1, \varkappa_2, \varkappa_3)$ is a triplet of integers with the values $\varkappa_\nu = 0,1,2 \ldots G_\nu - 1$. In total, there are hence $G_1 G_2 G_3$ such triplets to be considered.

For many purposes, it is now convenient to introduce the primitive translations of the reciprocal lattice, $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$, satisfying the relations $\mathbf{a}_k \cdot \mathbf{b}_l = \delta_{kl}$. To the triplet $(\varkappa_1, \varkappa_2, \varkappa_3)$, we will now associate the following vector $\mathbf{k}$ in the reciprocal lattice

$$\mathbf{k} = \frac{\varkappa_1}{G_1} \mathbf{b}_1 + \frac{\varkappa_2}{G_2} \mathbf{b}_2 + \frac{\varkappa_3}{G_3} \mathbf{b}_3 \tag{173}$$

For the inner product with $\mathfrak{m} = \mu_1 \mathfrak{a}_1 + \mu_2 \mathfrak{a}_2 + \mu_3 \mathfrak{a}_3$ one obtains

$$\mathfrak{k} \cdot \mathfrak{m} = \frac{\kappa_1}{G_1}\mu_1 + \frac{\kappa_2}{G_2}\mu_2 + \frac{\kappa_3}{G_3}\mu_3 , \tag{174}$$

and one can now write (172) in the form

$$\boxed{\mathbb{O}(\mathfrak{k}) = (G_1 G_2 G_3)^{-1} \sum_{(\mathfrak{m})} e^{-2\pi i \mathfrak{k}\cdot\mathfrak{m}} \, T(\mathfrak{m})} \tag{175}$$

According to the general theory, this operator fulfils the following basic relations:

$$\{ \mathbb{O}(\mathfrak{k}) \}^2 = \mathbb{O}(\mathfrak{k}) \quad , \qquad \mathbb{O}(\mathfrak{k})\,\mathbb{O}(\mathfrak{l}) = 0 , \; \mathfrak{k} \neq \mathfrak{l} \tag{176}$$

$$T(\mathfrak{m})\,\mathbb{O}(\mathfrak{k}) = e^{2\pi i \mathfrak{k}\cdot\mathfrak{m}}\,\mathbb{O}(\mathfrak{k}) , \tag{177}$$

$$I = \sum_{\mathfrak{k}} \mathbb{O}(\mathfrak{k}) \tag{178}$$

Every element Φ of our linear space satisfying (170) may hence be resolved into $G_1 G_2 G_3$ components

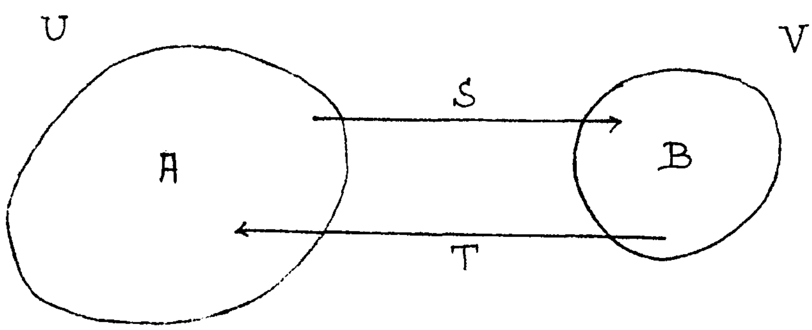
$$\Phi(\mathfrak{r}) = \sum_{\mathfrak{k}} \mathbb{O}(\mathfrak{k})\,\Phi(\mathfrak{r}) = \sum_{\mathfrak{k}} \Phi(\mathfrak{k},\mathfrak{r}) , \tag{179}$$

where $\Phi(\mathfrak{k},\mathfrak{r})$ may be characterized as the "Bloch components" of Φ

This approach may be used as a starting point for crystal theory and, for further details, we will refer elsewhere [13].

8. LINEAR MAPPING OF ONE LINEAR SPACE ON ANOTHER; MIRROR THEOREM

In this section, we will return to the study of finite linear spaces, and we will now consider two spaces U and V of order m and n, respectively, having the elements A and B. We will further consider two linear mappings, S and T, of which the first maps U on V and the second V on U. They correspond hence to linear operators which transform one linear space into another. Previously we have only considered operators which map a linear space on itself or on a subspace of itself.



Let us span the space U by a basis $\mathbf{X} = (X_1, X_2, \ldots X_m)$ and the space V by a basis $\mathbf{Y} = (Y_1, Y_2, \ldots Y_n)$. Since SX_k is an element of V and TY_ℓ an element of U, one obtains the unique expansions

$$
\begin{aligned}
S X_k &= \sum_{\ell=1}^{n} Y_\ell S_{\ell k}, && k = 1, 2, \ldots m \\
T Y_\ell &= \sum_{k=1}^{m} X_k T_{k\ell}, && \ell = 1, 2, \ldots n
\end{aligned}
\tag{180}
$$

where the coefficients $S_{\ell k}$ and $T_{k\ell}$ form rectangular matrices:

$$
\mathbf{S}_Y = {}^{n}\overset{m}{\boxed{\quad\quad\quad}}\,, \qquad \mathbf{T}_X = \overset{n}{\boxed{\;\;}}\,{}^{m}
\tag{181}
$$

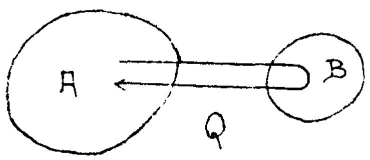
which are said to be the matrix representations of S and T with respect to the bases involved. For arbitrary elements $A = \mathbf{X}\mathbf{a}$ and $B = \mathbf{Y}\mathbf{b}$, one obtains directly

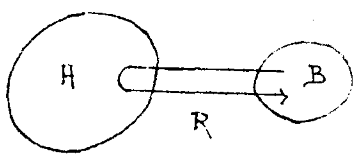
$$SA = \mathbf{Y}\,\mathbf{S}_Y\,\mathbf{a}, \qquad TB = \mathbf{X}\,\mathbf{T}_X\,\mathbf{b} \tag{182}$$

Let us now consider also the double mappings

$$Q = TS, \qquad R = ST, \tag{183}$$

which are illustrated by the figures below.





This implies that Q maps A on itself (or on a subspace of itself), whereas R maps B on itself. The operator Q may be represented by a quadratic matrix of order m in the basis $\mathbf{X}$, whereas R is represented by a quadratic matrix of order n in the basis $\mathbf{Y}$. From (180), one obtains directly

$$\mathbf{T}\cdot\mathbf{S} = \mathbf{Q}, \qquad \mathbf{S}\cdot\mathbf{T} = \mathbf{R} \tag{184}$$

Since the two operators R and Q are of different orders, it does not seem likely that they should be closely related. However, one has the fundamental theorem:

> The non-vanishing eigenvalues of ST and TS are identical, even with respect to their multiplicity. (185)

The theorem implies that, if $m > n$, there are at least $(m-n)$ eigenvalues of R which are vanishing. The proof can be based on the concept of the trace since, according to (57), one has

$$
\begin{aligned}
\mathrm{Tr}(Q) &= \mathrm{Tr}(TS) = \mathrm{Tr}(ST) = \mathrm{Tr}(R), \\
\mathrm{Tr}(Q^2) &= \mathrm{Tr}(TSTS) = \mathrm{Tr}(STST) = \mathrm{Tr}(R^2), \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{etc.}
\end{aligned}
\tag{186}
$$

which easily proves the conclusion. However, here we will also proceed in another way which gives us some other aspects on the problem.

Let us denote the eigenelements and eigenvalues of Q by u_k and a_k, respectively, and the corresponding quantities for R by v_ℓ and b_ℓ, so that

$$
Q u_k = a_k u_k \quad , \qquad R v_\ell = b_\ell v_\ell . \tag{187}
$$

The operation S maps further u_k on the element $\tilde{v}_k$, which in turn is mapped on $\tilde{\tilde{u}}_k$ by the operation T. Similarly T maps v_ℓ on $\tilde{u}_\ell$, which is then mapped on $\tilde{\tilde{v}}_\ell$ by S. Hence we have

$$
\begin{aligned}
S u_k &= \tilde{v}_k , \qquad & T \tilde{v}_k &= \tilde{\tilde{u}}_k , \\
T v_\ell &= \tilde{u}_\ell , \qquad & S \tilde{u}_\ell &= \tilde{\tilde{v}}_\ell .
\end{aligned}
\tag{188}
$$

Using (183) and (187), we obtain $\tilde{\tilde{u}}_k = T\tilde{v}_k = TSu_k = Qu_k = a_k u_k$ which shows that, if $a_k \neq 0$, one has $\tilde{\tilde{u}}_k \neq \bar{0}$ and consequently also $\tilde{v}_k \neq \bar{0}$. Similarly, one has $\tilde{\tilde{v}}_\ell = S\tilde{u}_\ell = STv_\ell = Rv_\ell = b_\ell v_\ell$ which implies that, if $b_\ell \neq 0$, one has $\tilde{\tilde{v}}_\ell \neq \bar{0}$ and also $\tilde{u}_\ell \neq \bar{0}$. For eigenelements associated with non-vanishing eigenvalues, the two image elements considered are hence different from the zero-element.

Let us now consider the properties of $\tilde{v}_k$ in greater detail for $a_k \neq 0$. One obtains directly

$$
\begin{aligned}
R\tilde{v}_k &= ST\tilde{v}_k = STSu_k = SQu_k = \\
&= S(a_k u_k) = a_k \tilde{v}_k ,
\end{aligned}
\tag{189}
$$

which shows that a_k is also an eigenvalue to R, since $\tilde{v}_k \neq \bar{0}$. For $b_\ell \neq 0$, one obtains in the same way

$$Q\tilde{u}_\ell = TS\tilde{u}_\ell = TST v_\ell = TRv_\ell = \tag{190}$$
$$= T(b_\ell v_\ell) = b_\ell \tilde{u}_\ell ,$$

which shows that b_ℓ is also an eigenvalue to Q, since $\tilde{u}_\ell \neq \bar{0}$. The non-vanishing eigenvalues to $Q = TS$ and $R = ST$ are hence necessarily the same.

This completes the proof for the non-degenerate case. In the case of a finite degeneracy and $m = 1$, one simply spans the eigenspace by a linearly independent set of eigenelements u_k', u_k'', u_k''', . . . , and consideration of the associated image elements according to (188) shows the validity of the theorem (185). The case of $m \neq 1$ requires somewhat more care, and it will be left out of our present discussion.

Conjugation of elements. - Let us consider the non-vanishing eigenvalues $a_p = b_p \neq 0$, and let us arrange the eigenelements to Q and R in pairs, so that $\tilde{v}_p = a_p^{+1/2} v_p$ and $\tilde{u}_p = a_p^{+1/2} u_p$, or

$$\boxed{\begin{aligned} v_p &= a_p^{-1/2} S u_p ; \\ u_p &= a_p^{-1/2} T v_p . \end{aligned}} \tag{191}$$

We note that each one of the relation (191) follows from the other, and we say that the elements u_p and v_p form a conjugated pair. It is also essential that the relation (191) contains a "phase convention".

Adding also the eigenelements associated with $a_k = 0$, one obtains a linearly independent set $u_1, u_2, \ldots u_m$ which may be used as a basis for the space U , and similarly the set $v_1, v_2, \ldots v_n$ may be used as a basis for V. A comparison between (180) and (191) in the form $Su_p = a_p^{1/2} v_p$, $Tv_p = a_p^{1/2} u_p$, shows that S and T may be given the matrix representation

$$S = \left(\begin{array}{cccc|c} a_1^{1/2} & & & & \\ & a_2^{1/2} & & & \mathbf{0} \\ & & \ddots & & \\ & & & a_m^{1/2} & \end{array} \right) , \quad T = \left(\begin{array}{cccc} a_1^{1/2} & & & \\ & a_2^{1/2} & & \\ & & \ddots & \\ & & & a_m^{1/2} \\ \hline & & \mathbf{0} & \end{array} \right) \tag{192}$$

and these rectangular matrices are hence brought to a kind of "diagonal form".

Using the symbols introduced in (11) and (12), one can now express an arbitrary element A or B in the form

$$\mathsf{A} = \sum_k \mathsf{X}_k \{\mathsf{X}_k, \mathbf{X} \mid \mathsf{A}\}, \qquad (193)$$
$$\mathsf{B} = \sum_l \mathsf{Y}_l \{\mathsf{Y}_l, \mathbf{Y} \mid \mathsf{B}\}.$$

Choosing $\mathbf{X} = \mathbf{u}$ and $\mathbf{Y} = \mathbf{v}$, this gives the "eigenexpansions"

$$\mathsf{A} = \sum_k u_k \{u_k, \mathbf{u} \mid \mathsf{A}\}, \qquad (194)$$
$$\mathsf{B} = \sum_l v_l \{v_l, \mathbf{v} \mid \mathsf{B}\}.$$

Letting Q work on the expression for A, one obtains $Q\mathsf{A} = \sum_k a_k u_k \{u_k, \mathbf{u} \mid \mathsf{A}\}$, where only the eigenelements having $a_k \neq 0$ will contribute. In this way, one obtains the symbolic relations

$$Q = \sum_p a_p u_p \{u_p, \mathbf{u} | , \qquad (195)$$
$$R = \sum_p a_p v_p \{v_p, \mathbf{v} | .$$

Letting S work on the expression for A in (194) and using the "conjugation" relations (191), one obtains $S\mathsf{A} = \sum_k a_k^{1/2} v_k \{u_k, \mathbf{u} \mid \mathsf{A}\}$, or symbolically

$$\boxed{\begin{aligned} S &= \sum_p a_p^{1/2} v_p \{u_p, \mathbf{u} | , \\ T &= \sum_p a_p^{1/2} u_p \{v_p, \mathbf{v} | . \end{aligned}} \qquad (196)$$

These relations may be considered as some form of "spectral resolution" for the operators S and T which map one linear space on another. Since such a mapping is quite common in quantum theory, the "mirror theorem" (185) and the associated relations are of fundamental importance in this connection. Of particular importance are the applications to density matrices [14] and to spin pairings [15].

7. INTRODUCTION OF A SCALAR PRODUCT; FROM LINEAR ALGEBRA TO VECTOR ALGEBRA.

Definitions. - Let us start by considering a linear space V of finite order n having a basis $\mathbf{X} = (X_1, X_2, \ldots X_n)$. According to (5), one has an expansion theorem in which the coefficients are uniquely determined and, using the symbol (12), one can write

$$A = \sum_k X_k \{ X_k, \mathbf{X} \mid A \} . \tag{197}$$

Here the notation $\{ X_k, \mathbf{X} \mid A \}$ simply means the coefficients for X_k in an expansion in terms of the basis $\mathbf{X}$ of the specific element A . Considering the expansions for $(A_1 + A_2)$ and αA , one obtains the relations

$$\begin{aligned} \{ X_k, \mathbf{X} \mid A_1 + A_2 \} &= \{ X_k, \mathbf{X} \mid A_1 \} + \{ X_k, \mathbf{X} \mid A_2 \} , \\ \{ X_k, \mathbf{X} \mid \alpha A \} &= \alpha \{ X_k, \mathbf{X} \mid A \} , \end{aligned} \tag{198}$$

showing the linear character of the symbol $\{ \}$. The definitions give further

$$\{ X_k, \mathbf{X} \mid X_l \} = \delta_{kl} . \tag{199}$$

Let us now introduce the concept of the scalar product . To every pair of elements, A and B, of the linear space we will associate a complex number called the scalar product and denoted by the symbol $\langle A | B \rangle$, which satisfies the following axioms :

(1) $\langle A | B_1 + B_2 \rangle = \langle A | B_1 \rangle + \langle A | B_2 \rangle$,

$\langle A | \alpha B \rangle = \alpha \langle A | B \rangle$;

$$(2) \qquad \langle A | B \rangle = \langle B | A \rangle^* \,,$$

$$(3) \qquad \langle A | A \rangle \geq 0 \,, \quad \text{and} \qquad (200)$$

$$\langle A | A \rangle = 0 \,, \quad \text{if and only if} \quad A = \bar{0}$$

The axiom (1) is essentially of the same type as (198), whereas (2) and (3) contain new properties which we have not used before in our treatment. The quantity $\langle A | A \rangle^{1/2}$ is often called the "length" of A and is denoted by $\| A \|$. We note that, even if the scalar product is given more properties than the symbol $\{ \}$, the axioms in (200) do not contain any recipe for the evaluation of this quantity, and there may actually exist many "realizations" of the scalar product. The vector algebra obtained from the linear algebra by adding the concept of the scalar product has hence an abstract but also very general character.

In connection with the notations, we observe that the bracket $\langle A | B \rangle$ is a physicist's symbol and that the mathematicians denote the same quantity by (B, A). According to (200), one obtains particularly $\langle \alpha A | B \rangle = \alpha^* \langle A | B \rangle$ showing the conjugate complex character associated with the first position. Two elements A and B are finally said to be orthogonal, if $\langle A | B \rangle = 0$. From the axioms (200), one can derive some important inequalities. If λ is a real parameter, one has

$$\langle A + \lambda B | A + \lambda B \rangle \geq 0 \,,$$

$$\langle A | A \rangle + \lambda \{ \langle A | B \rangle + \langle B | A \rangle \} + \lambda^2 \langle B | B \rangle \geq 0 \,. \qquad (201)$$

This implies that the discriminant can never be positive, i.e.

$$| \mathrm{Re} \{ \langle A | B \rangle \} |^2 - \langle A | A \rangle \langle B | B \rangle \leq 0 \,, \qquad (202)$$

where $\mathrm{Re} \{ \langle A | B \rangle \} = \frac{1}{2} [\langle A | B \rangle + \langle B | A \rangle]$ is the real part of the scalar product $\langle A|B\rangle$. This relation is true even if B is replaced by $e^{i\alpha}B$ and, by a convenient choice of α, one obtains

$$|\langle A|B\rangle|^2 \leq \langle A|A\rangle\langle B|B\rangle . \tag{203}$$

This is the famous **Schwarz's** inequality. **Using** this inequality for the cross-terms in $\langle A+B|A+B\rangle$, one **gets further** the "triangular inequality"

$$\big|\, \|A\| - \|B\| \,\big| \leq \|A+B\| \leq \|A\| + \|B\| . \tag{204}$$

Such relations are, of course, of essential importance in studying upper and lower bounds, questions of convergence etc.

Expansion coefficients as scalar products. - Let us now return to the expansion theorem (197). From the elements of the basis $\mathbf{X} = (X_1, X_2, \ldots X_n)$, one can construct a total of n^2 scalar products

$$\Delta_{kl} = \langle X_k | X_l \rangle , \tag{205}$$

which together form a matrix $\mathbf{\Delta}$ called the "metric" matrix. It is easily shown that the set $\mathbf{X}$ is linearly independent, if and only if $\det\{\mathbf{\Delta}\} \neq 0$. For a basis, the inverse matrix $\mathbf{\Delta}^{-1}$ will hence exist. From the expansion theorem $A = \sum_l X_l a_l = \mathbf{X}\mathbf{a}$, one obtains directly

$$\begin{aligned} \langle X_k | A \rangle &= \langle X_k | \sum_l X_l a_l \rangle = \sum_l \langle X_k | X_l \rangle a_l = \\ &= \sum_l \Delta_{kl} a_l , \end{aligned} \tag{206}$$

or, in matrix form, $\langle \mathbf{X} | A \rangle = \mathbf{\Delta}\mathbf{a}$, which gives

$$\boxed{\mathbf{a} = \mathbf{\Delta}^{-1} \langle \mathbf{X} | A \rangle .} \tag{207}$$

This formula gives the expansion coefficients expressed in terms of scalar products, but it does not give any recipe for the evaluation of these coefficients, unless one has a "realization" of the scalar product. For the components of (207), one obtains

$$a_k = \sum_l \Delta^{-1}_{kl} \langle X_l | A \rangle . \tag{208}$$

The $\{ \ \}$- symbol in (197) may then be expressed in the form

$$\{ X_k, X | = \sum_l \Delta^{-1}_{kl} \langle X_l | , \tag{209}$$

and, for the projection operator O_k in (52), one has

$$O_k = X_k \{ X_k, X | = \sum_l X_k \Delta^{-1}_{kl} \langle X_l | . \tag{210}$$

For the resolution of the identity (49), this gives particularly

$$I = \sum_{k,l} X_k \Delta^{-1}_{kl} \langle X_l | . \tag{211}$$

Of special importance is the case of an orthonormal basis satisfying the relations

$$\langle X_k | X_l \rangle = \delta_{kl} , \tag{212}$$

or $\Delta = 1$. By means of Schmidt's successive orthogonalization procedure, it is easily seen that, by a convenient linear transformation, every basis may be brought to orthonormal form. In this case, the previous formulas may be simplified to the form

$$a_k = \langle X_k | A \rangle ; \tag{213}$$

$$O_k = X_k \langle X_k | ; \qquad I = \sum_k X_k \langle X_k | . \tag{214}$$

Let us now return to the case of a general metric matrix $\mathbf{\Delta}$ and consider the matrix representation of a linear operator T defined by (31) or (32). According to (209), one obtains immediately

$$T_{k\ell} = \{ X_k, \mathbf{X} \mid TX_\ell \} = \sum_\alpha \Delta^{-1}_{k\alpha} \langle X_\alpha \mid TX_\ell \rangle \tag{215}$$

By denoting the matrix formed by the scalar products $\langle X_k | TX_\ell \rangle \equiv \langle X_k | T | X_\ell \rangle$ by $\mathfrak{T}$, one gets hence

$$\boxed{\mathbf{T} = \mathbf{\Delta}^{-1} \mathfrak{T}} \tag{216}$$

In quantum theory, $\mathfrak{T}$ is very often described as the matrix of T with respect to the basis $\mathbf{X}$, whereas, in linear algebra, this name refers to the matrix $\mathbf{T}$. Note particularly that, for an operator product FG , one has, according to (39) and (216) :

$$(FG)_{\mathbf{X}} = \mathbf{\Delta}^{-1} \mathfrak{F} \mathbf{\Delta}^{-1} \mathfrak{G} , \tag{217}$$

where $\mathfrak{F}$ and $\mathfrak{G}$ are the matrices formed by the elements $\langle X_k | F | X_\ell \rangle$ and $\langle X_k | G | X_\ell \rangle$, respectively.

It is now also easy to understand the connection between the simple form (65) of the eigenvalue problem in matrix representation in linear algebra and the conventional form in quantum mechanics. From (65) and (216) follows

$$\begin{aligned} (\mathbf{T} - \lambda \cdot \mathbf{1}) \mathbf{c} &= 0 , \\ (\mathbf{\Delta}^{-1} \mathfrak{T} - \lambda \cdot \mathbf{1}) \mathbf{c} &= 0 , \\ (\mathfrak{T} - \lambda \cdot \mathbf{\Delta}) \mathbf{c} &= 0 . \end{aligned} \tag{218}$$

or

$$\sum_{l} \{ \langle X_k | T | X_l \rangle - \lambda \langle X_k | X_l \rangle \} c_l = 0 . \quad (219)$$

The last form is well-known from the applications to e.g. quantum chemistry.

Projection on a linear manifold. - In connection with the general expansion theorem, it is convenient to study also the concept of the projection of an arbitrary element A onto a linear manifold imbedded in the space. Let the linear manifold be spanned by the elements $\mathbf{f} = (f_1, f_2, \ldots f_m)$, and let us determine the coefficients a_k in the expansion

$$A = \sum_{k=1}^{m} f_k a_k + R , \quad (220)$$

so that the length of the remainder element R becomes as small as possible, so that

$$\| R \|^2 = \langle A - \mathbf{f}\mathbf{a} | A - \mathbf{f}\mathbf{a} \rangle = \text{minimum} . \quad (221)$$

For this purpose, we will introduce the metrix matrix $\mathbf{\Delta} = \langle \mathbf{f} | \mathbf{f} \rangle$ and the vector $\mathbf{c} = \mathbf{\Delta}^{-1} \langle \mathbf{f} | A \rangle$ according to (207), which gives $\langle \mathbf{f} | A \rangle = \mathbf{\Delta}\, \mathbf{c} = \langle \mathbf{f} | \mathbf{f}\mathbf{c} \rangle$ and $\langle A | \mathbf{f} \rangle = \langle \mathbf{f}\mathbf{c} | \mathbf{f} \rangle$. Using the axioms (200), one hence obtains the identity

$$\begin{aligned} &\langle A - \mathbf{f}\mathbf{a} | A - \mathbf{f}\mathbf{a} \rangle = \\ &= \langle A | A \rangle - \langle \mathbf{f}\mathbf{a} | A \rangle - \langle A | \mathbf{f}\mathbf{a} \rangle + \langle \mathbf{f}\mathbf{a} | \mathbf{f}\mathbf{a} \rangle = \\ &= \langle A | A \rangle - \langle \mathbf{f}\mathbf{a} | \mathbf{f}\mathbf{c} \rangle - \langle \mathbf{f}\mathbf{c} | \mathbf{f}\mathbf{a} \rangle + \langle \mathbf{f}\mathbf{a} | \mathbf{f}\mathbf{a} \rangle = \\ &= \langle A | A \rangle - \langle \mathbf{f}\mathbf{c} | \mathbf{f}\mathbf{c} \rangle + \langle \mathbf{f}(\mathbf{c}-\mathbf{a}) | \mathbf{f}(\mathbf{c}-\mathbf{a}) \rangle \geq 0 \end{aligned} \quad (222)$$

Only the last term contains the coefficients $\mathbf{a}$, and we note that it can never be negative and has its minimum for $\mathbf{a} = \mathbf{c} = \mathbf{\Delta}^{-1} \langle \mathbf{f} | A \rangle$

For $\mathbf{a} = \mathbf{c}$, one obtains particularly

$$\langle A - \mathbf{f}\mathbf{c} \mid A - \mathbf{f}\mathbf{c} \rangle = \langle A \mid A \rangle - \langle \mathbf{f}\mathbf{c} \mid \mathbf{f}\mathbf{c} \rangle \geq 0 . \tag{223}$$

If $\mathbf{c}^{\dagger}$ denotes the row vector of the elements $(c_1^*, c_2^*, \ldots c_m^*)$, one has hence $\langle \mathbf{f}\mathbf{c} | \mathbf{f}\mathbf{c} \rangle = \mathbf{c}^{\dagger} \langle \mathbf{f} | \mathbf{f} \rangle \mathbf{c} = \mathbf{c}^{\dagger} \mathbf{\Delta}\, \mathbf{c}$ and the inequality

$$\langle A \mid A \rangle \geq \sum_{k,l=1}^{m} c_k^* \Delta_{kl} c_l , \tag{224}$$

which is a generalization of Bessel's inequality to the case of an arbitrary metric.

According to (220), one can now write

$$A = \mathbf{f}\mathbf{c} + R , \tag{225}$$

where the term $\mathbf{f}\mathbf{c}$ is called the "projection" of A on the linear manifold spanned by $\mathbf{f} = (f_1, f_2, \ldots f_m)$. One obtains

$$\mathbf{f}\mathbf{c} = \mathbf{f} \mathbf{\Delta}^{-1} \langle \mathbf{f} | A \rangle = \mathbb{O} A , \tag{226}$$

where

$$\boxed{\mathbb{O} = \mathbf{f} \mathbf{\Delta}^{-1} \langle \mathbf{f} | ,} \tag{227}$$

is said to be the projection operator associated with the manifold $\mathbf{f}$. One gets directly $O^2 = O$. Since further

$$\begin{aligned} \langle R \mid \mathbf{f}\mathbf{c} \rangle &= \langle A - \mathbf{f}\mathbf{c} \mid \mathbf{f}\mathbf{c} \rangle = \\ &= \langle A \mid \mathbf{f}\mathbf{c} \rangle - \langle \mathbf{f}\mathbf{c} \mid \mathbf{f}\mathbf{c} \rangle = \\ &= \langle \mathbf{f}\mathbf{c} \mid \mathbf{f}\mathbf{c} \rangle - \langle \mathbf{f}\mathbf{c} \mid \mathbf{f}\mathbf{c} \rangle = 0 , \end{aligned} \tag{228}$$

the remainder element R is orthogonal to the projection $\mathbf{f}\mathbf{c}$.

Let us now consider an infinite space and let us assume that any finite subset of the set $(f_1, f_2, f_3, \ldots)$ is linearly independent. For every value of m, one has, according to (224),

$$\langle A | A \rangle \geq \sum_{k,l=1}^{m} c_k^* \Delta_{kl} c_l , \tag{229}$$

where it is easily proven that the right-hand sum for $m = 1, 2, 3 \ldots$ forms a series of never decreasing positive numbers. Since all the partial sums have an upper bound, the limit for $m \longrightarrow \infty$ exists, and one has

$$\lim_{m \to \infty} \sum_{k,l=1}^{m} c_k^* \Delta_{kl} c_l \leq \langle A | A \rangle . \tag{230}$$

If the equality sign is valid, one says that the infinite set $(f_1, f_2, f_3 \ldots)$ is complete, and one has obtained a generalization of Parseval's relation. For the case of an orthonormal basis, $\boldsymbol{\Delta} = \mathbf{1}$, one obtains the conventional form :

$$\langle A | A \rangle = \sum_{k=1}^{\infty} | c_k |^2 \tag{231}$$

Returning to an arbitrary metric, we note that, according to (227), one can also write (230) in the form

$$\lim_{m \to \infty} \langle A | I - \mathbb{O}_m | A \rangle = 0 . \tag{232}$$

The expansion theorem, on the other hand, takes the form

$$\lim_{m \to \infty} (I - \mathbb{O}_m) A = \bar{0} , \tag{233}$$

and we note that, in general, there is a considerable difference in convergence properties between (232) and (233), and that one relation does not necessarily follow from the other. The property (232) is called "convergence in mean" .

Hilbert space. - In order to be able to discuss convergence properties in general, one has often introduced an additional axiom which leads to the concept of the "Hilbert space" . Such a space is an infinite vector space which contains also its limiting elements :

> If $A_1, A_2, A_3, \ldots$ is a set of elements in the space having the property
>
> $$\| A_m - A_{m+n} \| < \varepsilon , \tag{234}$$
>
> as soon as $m > m(\varepsilon)$, then there exists an element A in the space such that $\| A_m - A \| < \varepsilon$, and one writes
>
> $$\lim_{m \to \infty} A_m = A .$$

In addition, one introduces also a "separability axiom" stating that every element A may be reached by a denumerably infinite set of elements $A_1, A_2, A_3, \ldots$ such that $\| A_m - A \| < \varepsilon$, as soon as $m > m(\varepsilon)$. For a detailed treatment of the properties of the Hilbert space, we will refer to the excellent books available [16)]. It should be observed that the terminology introduced in connection with the linear algebra and particularly the concepts of projection operators, resolution of the identity, and spectral resolution of an operator play an important role also in this connection.

Pair of adjoint operators; normal and self-adjoint operators. - In conclusion, we will briefly survey some of the fundamental concepts as to linear operators which are introduced on the basis of the scalar product. Let T be an arbitrary operator having the domain D_T . Two operators T and $T^\dagger$ are said to form a pair of adjoint operators, if they have the same domains and further

$$\langle A | T^\dagger | A \rangle = \langle A | T | A \rangle^* , \tag{235}$$

for every element A in the domain. From the definition follows the theorem :

$$\langle TB | A \rangle = \langle B | T^{\dagger} A \rangle , \tag{236}$$

provided A and B belong to D_T. In order to prove this "turn-over rule", one uses the following identity :

$$\begin{aligned} \langle B | T^{\dagger} A \rangle &\equiv \frac{1}{4} \{ \langle B+A | T^{\dagger} | B+A \rangle - \langle B-A | T^{\dagger} | B-A \rangle - \\ &\quad - i \langle B+iA | T^{\dagger} | B+iA \rangle + i \langle B-iA | T^{\dagger} | B-iA \rangle \} = \\ &= \frac{1}{4} \{ \langle B+A | T | B+A \rangle - \langle B-A | T | B-A \rangle + \\ &\quad + i \langle B+iA | T | B+iA \rangle - i \langle B-iA | T | B-iA \rangle \}^{*} = \\ &= \langle A | T | B \rangle^{*} = \langle TB | A \rangle , \end{aligned} \tag{237}$$

which completes the proof. Using (236), one obtains the well-known rules $(F+G)^{\dagger} = F^{\dagger} + G^{\dagger}$, $(FG)^{\dagger} = G^{\dagger} F^{\dagger}$./An operator Λ is said to be <u>normal</u>, if it commutes with its adjoint operator $\Lambda^{\dagger}$, so that

$$\Lambda \Lambda^{\dagger} = \Lambda^{\dagger} \Lambda . \tag{238}$$

If the operator Λ has the eigenelement Φ_k associated with the eigenvalue λ_k, the operator $\Lambda^{\dagger}$ has the same eigenelement associated with the eigenvalue λ_k^{*}, so that

$$\Lambda \Phi_k = \lambda_k \Phi_k , \qquad \Lambda^{\dagger} \Phi_k = \lambda_k^{*} \Phi_k . \tag{239}$$

The proof follows from the fact that

$$\begin{aligned} \| (\Lambda^{\dagger} - \lambda_k^{*}) \Phi_k \|^2 &= \langle (\Lambda^{\dagger} - \lambda_k^{\dagger}) \Phi_k | (\Lambda^{\dagger} - \lambda_k^{\dagger}) \Phi_k \rangle = \\ &= \langle \Phi_k | (\Lambda - \lambda_k)(\Lambda^{\dagger} - \lambda_k^{\dagger}) \Phi_k \rangle = \\ &= \langle \Phi_k | (\Lambda^{\dagger} - \lambda_k^{\dagger})(\Lambda - \lambda_k) \Phi_k \rangle = \\ &= \langle (\Lambda - \lambda_k) \Phi_k | (\Lambda - \lambda_k) \Phi_k \rangle = \\ &= \| (\Lambda - \lambda_k) \Phi_k \|^2 = 0 . \end{aligned} \tag{240}$$

The normal operators are characterized by the fact that eigenelements Φ_k and Φ_l associated with different eigenvalues, $\lambda_k \neq \lambda_l$, are necessarily orthogonal :

$$\langle \Phi_k | \Phi_l \rangle = 0 , \qquad \lambda_k \neq \lambda_l . \tag{241}$$

One has $\lambda_k \langle \Phi_k | \Phi_l \rangle = \langle \lambda_k^* \Phi_k | \Phi_l \rangle = \langle \Lambda^\dagger \Phi_k | \Phi_l \rangle =$
$= \langle \Phi_k | \Lambda \Phi_l \rangle = \langle \Phi_k | \lambda_l \Phi_l \rangle = \lambda_l \langle \Phi_k | \Phi_l \rangle$, i.e.
$(\lambda_k - \lambda_l) \langle \Phi_k | \Phi_l \rangle = 0$, which proves the theorem.

For a finite space, a normal operator may, of course, be brought to classical canonical form and, using (238), one can show that this form must necessarily be diagonal. For the largest Segré characteristic m_k to each eigenvalue, one obtains $m_k = 1$, and the reduced Cayley-Hamilton equation is then of the type (148). The property of normality is hence of essential importance in the projection operator approach.

Using the eigenvalue relation (239), one easily obtains

$$\lambda_k = \frac{\langle \Phi_k | \Lambda | \Phi_k \rangle}{\langle \Phi_k | \Phi_k \rangle} \tag{242}$$

The corresponding quantity for an arbitrary element A :

$$\langle \Lambda \rangle_{Av} = \frac{\langle A | \Lambda | A \rangle}{\langle A | A \rangle} , \tag{243}$$

is called the "expectation value" of Λ with respect to A . For elements close to the eigenelements, there is an important "variation principle" . Putting $A = \Phi_k + \delta\Phi_k$ and using the relation $(\Lambda - \lambda_k I)\Phi_k = \bar{0}$ to the left and to the right, one gets

$$\langle \Lambda - \lambda_k \rangle_{Av} = \frac{\langle A | \Lambda - \lambda_k | A \rangle}{\langle A | A \rangle} = \frac{\langle \delta\Phi_k | \Lambda - \lambda_k | \delta\Phi_k \rangle}{\langle A | A \rangle} ,$$

i.e.

$$\langle \Lambda \rangle_{Av} = \lambda_k + \frac{\langle \delta \Phi_k | \Lambda - \lambda_k | \delta \Phi_k \rangle}{\langle A | A \rangle} \tag{244}$$

For a normal operator, a first-order variation in Φ_k leads hence to a second-order variation in $\langle \Lambda \rangle_{Av}$, i.e. $\delta \langle \Lambda \rangle_{Av} = 0$. This property is of fundamental importance in the quantum-mechanical applications.

A special class of normal operators are the unitary operators U characterized by the relations

$$U^\dagger U = U U^\dagger = 1 \tag{245}$$

Their eigenvalues satisfy the relation $\lambda_k^* \lambda_k = 1$ and are thus situated on the unit circle in the complex plane.

Of particular importance are finally the self-adjoint or hermitean operators F which satisfy the relation $F^\dagger = F$, i.e.

$$\langle A | F | A \rangle = \text{real quantity} \tag{246}$$

According to (242), the eigenvalues are then real numbers. The self-adjoint operators are hence normal operators having their eigenvalues on the real axis. In quantum theory, all physical quantities are represented by operators having real expectation values, i.e. by self-adjoint operators, and they are hence of fundamental importance in the applications.

Additional pages to TN 125 "Linear Algebra and the Fundaments of Quantum Theory" by Per Olov Löwdin

Width of an operator; uncertaint relations. - Let T be an arbitrary linear operator having the domain D_T. If A is an element within this domain normalized to unity so that $\|A\| = 1$, one defines in accordance with (243) the expectation value T of the operator T with respect to A by the formula

$$\overline{T} = \langle T \rangle_{Av} = \langle A|T|A \rangle \tag{246a}$$

which is in general a complex number. The "width" ΔT of the operator T with respect to A is further defined by the relation

$$\Delta T = \| (T - \overline{T})A \| \tag{246b}$$

From the third axiom in (200) for the scalar product follows that the width ΔT vanishes, if and only if A is an eigenelement of T. This implies also that the width ΔT in a certain sense must be a measure of the deviation of A from an eigenelement. Using the definition, one obtains immediately the following transformation

$$\begin{aligned}(\Delta T)^2 &= \langle (T-\overline{T})A|(T - \overline{T})A \rangle = \\ &= \langle TA|TA \rangle - |\langle A|T|A \rangle|^2 \geq 0 .\end{aligned} \tag{246c}$$

If, in addition to A, the element TA is also situated within the domain D_T, one may apply the turn-over rule (237) which gives

$$\begin{aligned}(\Delta T)^2 &= \langle A|T^{\dagger}T|A \rangle - |\langle A|T|A \rangle|^2 = \\ &= \overline{T^{\dagger}T} - |\overline{T}|^2\end{aligned} \tag{246d}$$

We note that this formula is valid only for elements A within the domain of the operator $T^{\dagger}T$, which means that it is much more restricted than (246c).

Let us now consider a second linear operator R with the domain D_R, and let further A be a normalized element within the intersection of D_R and D_T. According to (246b) one has the definitions

$$\Delta T = \|(T - \overline{T})A\| \;;\quad \Delta R = \| (R - \overline{R}) A\| \;. \tag{246c}$$

Using Schwarz's inequality (203), one obtains the following transformation

$$\begin{aligned} \Delta T . \Delta R &= \| (T - \overline{T})A\| . \|(R-\overline{R})A \| \geq \\ &\geq |\langle (T - \overline{T})A| (R-\overline{R})A\rangle | \geq \\ &\geq | I \{\langle (T-\overline{T})A|(R-\overline{R})A\rangle\} | = \\ &= \frac{1}{2} | \langle (T -\overline{T})A| (R-\overline{R})A\rangle - \langle (T-\overline{T})A|(R-\overline{R})A\rangle^{*} | = \\ &= \frac{1}{2} |\langle (T- \overline{T})A| (R-\overline{R})A\rangle - \langle (R-\overline{R})A|(T-\overline{T})A\rangle | = \\ &= \frac{1}{2}|\langle TA|RA\rangle - \langle RA|TA\rangle - \overline{T}^{*}\overline{R} + \overline{R}^{*}\overline{T} |. \end{aligned} \tag{246f}$$

which is the <u>uncertainty relation</u> for a general pair of linear operators. By using the turn-over rule (237), one gets the much more restricted formula

$$\Delta T . \Delta R \geq \frac{1}{2} |\langle A|T^{\dagger}R - R^{\dagger}T|A\rangle - (\overline{T}^{*}\overline{R} - \overline{R}^{*}\overline{T})| \;. \tag{246g}$$

For a self-adjoint operator F, one has $\Delta F = \|(F - \overline{F})A\|$, whereas the special form (246d) gives the relation

$$\begin{aligned} (\Delta F)^2 &= \langle A|(F-\overline{F})^2 | A\rangle = \\ &= \overline{F^2} - \overline{F}^2 \;. \end{aligned} \tag{246h}$$

In using the statistical interpretation of quantum mechanics, the width ΔF is often described as the "quadratic deviation" of F from the average value $\overline{F}$. For a pair of self-adjoint operators F and G, the uncertainty relations (246f) takes the form

$$\Delta F . \Delta G \geq \frac{1}{2} | \langle FA|GA\rangle - \langle GA|FA\rangle | \;, \tag{246i}$$

for all elements A within the intersection of D_F and D_G. . Using (246g), one obtains the special form

$$\Delta F . \Delta G \geq \frac{1}{2}| \langle A|FG - GF| A\rangle | \;. \tag{246j}$$

restricted to the elements within the domain of the operator $(FG - GF)$.

This is the form of the uncertainty relations most well-known in the applications to quantum mechanics, and we note that it depends essentially only on Schwarz's inequality (203), i.e. on the axioms (200) for the scalar product. It is interesting to observe that the uncertainty relations are hence completely independent of any particular "realization" of the scalar product.

For the pair of self-adjoint operators $F = p = \frac{h}{2\pi i} \frac{d}{dx}$ and $G = x$, one has the commutation relation

$$px - xp = \frac{h}{2\pi i}, \tag{246k}$$

and application of (246j) leads to the special formula

$$\Delta p \cdot \Delta x \geq \frac{h}{4\pi}, \tag{246l}$$

which is Heisenberg's uncertainty relation for the position x and the momentum p. The more general form (246j) is due to Born.

8. DISCUSSION

The purpose of our study is to show that one can develop an appreciable part of the terminology and the conceptual framework associated with the fundaments of quantum theory by using only the axioms of the theory of linear spaces. The eigenvalue problem, the projection operators, the resolution of the identity, and the spectral resolution of an operator are concepts which may be reached and discussed in this way. The theorems for finite spaces are illustrative but are, of course, of an elementary nature. However, some of the theorems may be generalized also to infinite spaces.

In treating infinite spaces, we are considering only operators having all their eigenvalues situated in a finite number of points in the complex plane, each one of which has an infinite multiplicity. From the existence of a finite-order reduced Cayley-Hamilton equation, we have derived a set of projection operators which form a resolution of the identity and lead to a splitting of the space into a set of infinite subspaces associated with the eigenvalues. Every element A may then be uniquely resolved into components A_k which are eigenelements to the operator concerned. If there are several commuting operators, the procedure leads to a splitting of the space into simultaneous eigenspaces.

In quantum theory, this process is of particular importance in treating constants of motion. The Schrödinger equation is an eigenvalue problem of the form

$$\mathcal{H}\underline{\Psi} = E\underline{\Psi}, \tag{247}$$

and T is a constant of motion, if it commutes with H , so that

$$T\mathcal{H} = \mathcal{H}T . \tag{248}$$

Let T have the eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_m$ and the projection operators $O_1, O_2, \ldots O_n$ defined by (153) , so that $I = \sum_k O_k(T)$.

One has immediately

$$\Psi = \sum_k \Psi_k , \qquad \Psi_k = O_k \Psi , \tag{249}$$

which gives

$$\mathcal{H} \Psi_k = E \Psi_k , \qquad T \Psi_k = \lambda_k \Psi_k , \tag{250}$$

showing that the wave functions associated with a specific energy level E may be classified by means of the eigenvalues λ_k . Of still greater importance is this " component analysis" of an approximate eigenfunction[17].

Constants of motion which have been treated in this way include the spin [18], the various angular momenta in atomic theory [19], the general angular momenta [20], the exchange operators [21], and the translations [22]. In all these cases, one is considering a single operator Λ or a set of commuting operators $\Lambda_1, \Lambda_2, \ldots$ It is evident that, if one would have group of operators as constants of motion, one could utilize the well-known projection operators from the group algebra in exactly the same way for a splitting of the entire space. The theory of point groups would lead to new results, whereas the theory of continuous groups for translations and angular momenta would give essentially the results already obtained.

The component analysis is a tool which is of importance also in discussing the correlation problem associated with the one-particle model in physics and chemistry. In the Hartree-Fock scheme, the total wave function is approximated by a single determinant D , whereas, in the extended Hartree-Fock scheme, one has carried out a component analysis with respect to the constants of motion, so that

$$\Psi = O D , \tag{251}$$

where O is an appropriate projection operator selecting the component desired. In practical applications, this simple approach has given surprizingly good results [23].

A previous discussion of the constants of motion and their projection operators was based on the concept of the scalar product, but it is here shown that all the essential results can be obtained solely in the framework of linear algebra.

The introduction of the scalar product renders some further simplifications, for instance, in connection with the calculation of the expectation value of H with respect to the wave function (251) :

$$\langle \mathcal{H}_{op} \rangle_{Av} = \frac{\langle \mathcal{O}\mathcal{D} | \mathcal{H} | \mathcal{O}\mathcal{D} \rangle}{\langle \mathcal{O}\mathcal{D} | \mathcal{O}\mathcal{D} \rangle} = \frac{\langle \mathcal{D} | \mathcal{H}\mathcal{O} | \mathcal{D} \rangle}{\langle \mathcal{D} | \mathcal{O} | \mathcal{D} \rangle} , \qquad (252)$$

where we have used the formula $O^{\dagger}HO = OHO = HO^{2} = HO$. The component analysis is of particular importance in connection with the variation principle.

For a finite space, the eigenvalue problem of the type $TC = \lambda\ C$ is usually well-defined, but, for an infinite space, it may happen that some auxiliary <u>boundary conditions</u> are needed to determine the eigenvalue spectrum. The scalar product plays an important role in this connection, and, for the Schrödinger equation (247), one usually required that the solution Ψ should belong to the Hilbert space (closed states) or have a scalar product with the functions out of this space (scattering states).

The physical interpretations of quantum theory are finally based on the scalar product entering the "expectation value" (243) or

$$\langle F \rangle_{Av} = \frac{\langle \Psi | F | \Psi \rangle}{\langle \Psi | \Psi \rangle} . \qquad (253)$$

Introducing the eigenfunctions to F as a basis, this leads to the well-known probability interpretation of quantum theory. It has sometimes been said, that this interpretation depends on the existence of an expansion theorem, but it is, of course, sufficient that Parseval's relation (232) is fulfilled, i.e. that the system of eigenfunctions is complete.

A characteristic feature of the theory of linear algebra, vector algebra, and Hilbert space is that it can be developed in a very general form based solely on a system of axioms. This means that the theory itself does not give any explicit recipe for the evaluation of the quantities involved, and that there may exist many "realizations" of the abstract theory. Quantum theory is based on a specific recipe for evaluating the scalar product of e.g. the type

$$\langle A | B \rangle = \int A^*(x) B(x) (dx) \tag{254}$$

but the conceptual framework is independent of this particular realization.

The scalar product as a concept is certainly a very essential part of quantum theory which is usually introduced at the beginning in every theory. Here we have tried to see how far one could reach without this fundamental tool, and it turns out that a surprizingly large part of the conceptual framework is based on linear algebra alone.

REFERENCES

1. E. Schrödinger, Ann. Physik 79, 361 (1926).
2. W. Heisenberg, Zs. Physik 33, 879 (1925); M. Born and P. Jordan, Zs. Physik 34, 858 (1925); M. Born, W. Heisenberg, P. Jordan, Zs. Physik 35, 557 (1926).
3. P.A.M. Dirac, Proc. Roy. Soc. (London) A109, 642 (1925).
4. E. Schrödinger, Ann. Physik 79, 489 (1926).
5. J. von Neumann, "Math. Grundlagen der Quantenmechanik" (Springer, Berlin 1932).
6. P.A.M. Dirac, "The Principles of Quantum Mechanics" (Clarendon Press, Oxford, 1st ed. 1930, 4th ed. 1958).
7. P.O. Löwdin, Phys. Rev. 97, 1509 (1955); Colloq. Intern. centre Natl. recherche sci. (Paris) 82, 23 (1958); Technical Note 12, Uppsala Quantum Chemistry Group, Uppsala, Sweden (1958); Rev. Mod. Phys. 36, 966 (1964).
8. See e.g. the survey in P.O. Löwdin, J. Appl. Phys. Suppl. 33, 251 (1962).
9. P.O. Löwdin, Revs. Mod. Phys. 34, 520 (1962).
10. A.J. Coleman, "A Note on Nilpotent Operators, Jordan Canonical Form, Segré Characteristics, and Minimal Polynomials", Uppsala Quantum Chemistry Group (1962); unpublished.
11. See e.g. D.E. Littlewood, "Theory of Group Characters and Matrix Representations of Groups" (Clarendon Press, Oxford 1950).
12. F. Bloch, Z. Physik 52, 555 (1929); 57, 545 (1929).
13. P.O. Löwdin, Adv. Phys. 5, 1 (1956); see also reference 8.
14. E. Schmidt, Math. Ann. 63, 433 (1907); B.C. Carlson and J.M. Keller, Phys. Rev. 121, 659 (1961); A.J. Coleman, Revs. Mod. Phys. 35, 668 (1963).

15. A.T.Amos and G.G. Hall, Proc.Roy.Soc.(London) A263, 483 (1961); P.O.Löwdin, reference 8, particularly p.270.

16. See e.g. v. Neumann, ref.5; M.H.Stone, "Linear Transformations in Hilbert Space", (Amer.Math.Soc., New York 1932); F. Riesz and B. Sz.- Nagy, "Functional Analysis" (Frederick Ungar Publ.Co., New York 1955).

17. See reference 9.

18. See reference 7; H. McIntosh, J.Math.Phys. 1, 453 (1960); J.K.Percus and A. Rotenberg, J.Math.Phys. 3, 928 (1962); V.H.Smith, Jr., J.Chem.Phys. 41, 277 (1964).

19. R. Fieschi and P.O.Löwdin, Technical Note No.4, Uppsala Quantum Chemistry Group (1957; unpublished); P.O.Löwdin, Proc. Robert A. Welch Foundation, Conf.Chem.Res.II, Atomic Structure (1958); J.L.Calais and J. Linderberg, Technical Note No.39, Uppsala Quantum Chemistry Group (1960; to be published Theor.Chim.Acta 1964); A. Rotenberg, J.Chem.Phys. 39, 512 (1963).

20. See reference 7.

21. P.O.Löwdin, Revs.Mod.Phys. 34, 80 (1962).

22. See reference 13.

23. P.O.Löwdin, Nikko Symposium on Molecular Physics (Maruzen, Tokyo, 1954), p.13; Phys.Rev. 97, 1509 (1955); Proc. 10th Solvay Conference, 71 (1955); T. Itoh and H. Yoshizumi, J.Phys.Soc. Japan 10, 201 (1955); J.Chem.Phys. 23, 412 (1955); Busseiron Kenkyu 83, 13 (1955); R. Pauncz, J. de Heer, P.O.Löwdin, J.Chem.Phys. 36, 2247, 2257 (1962); J. de Heer, J.Chem. Phys. 37, 2080 (1962); J. de Heer, J.Phys.Chem. 66, 2288 (1962); R. Pauncz, J.Chem.Phys. 37, 2739 (1962); J. de Heer, Rev.Mod. Phys. 35, 631 (1963); this paper gives a survey of the results obtained so far; J. de Heer, R. Pauncz, J.Chem.Phys. 39, 2314 (1963); J.D.Swalen, J. de Heer, J.Chem.Phys. 40, 378 (1964); G.W.Parker, J.D. Memory, Bull.Am.Phys.Soc. 9, 102 (1964); J.L. Calais, Arkiv Fysik (1964), to be published; R.Pauncz, "Molecular Orbitals in Chemistry, Physics, and Biology" (Mulliken dedicatory volume; Ed. by P.O.Löwdin and B.Pullman, Academic Press, New York 1964) p.433.